

Kölner Arbeitspapiere zur Bibliotheks- und Informationswissenschaft

Band 76

Personal Digital Archiving

Melanie Engels

April 2015

Fachhochschule Köln

Fakultät für Informations- und Kommunikationswissenschaften

Institut für Informationswissenschaft

Entstanden als Masterthesis im Studiengang Bibliotheks- und Informationswissenschaft

Betreuer: Prof. Dr. Achim Oßwald

Engels, Melanie
Personal Digital Archiving.
Köln: Fachhochschule Köln,
Fakultät für Informations- und Kommunikationswissenschaften,
Institut für Informationswissenschaft, 2015
(Kölner Arbeitspapiere zur Bibliotheks- und Informationswissenschaft; 76)

ISSN (elektronische Version) 1434-1115

Die Kölner Arbeitspapiere zur Bibliotheks- und Informationswissenschaft berichten über aktuelle Forschungsergebnisse des Instituts Informationswissenschaft der Fachhochschule Köln. Veröffentlicht werden sowohl Arbeiten der Dozentinnen und Dozenten als auch herausragende Arbeiten der Studierenden. Die Kontrolle der wissenschaftlichen Qualität der Veröffentlichungen liegt bei der Schriftleitung. Jeder Band erscheint in elektronischer Version (über unsere Homepage: <http://www.fbi.fh-koeln.de/institut/papers/arbeitspapiere.php>).

Fachhochschule Köln
Fakultät für Informations- und Kommunikationswissenschaften
Institut für Informationswissenschaft
Claudiusstr.1 D-50678 Köln
Tel.: 0221/8275-3376, Fax: 0221/3318583

Schriftleitung: Prof. Dr. Hermann Rösch, Susanne Röltgen
© FH Köln 2015

ABSTRACT (DEUTSCH)

Personal Digital Archiving ist ein wenig untersuchtes Forschungsgebiet, das sich mit der Archivierung privater Daten durch private Anwender beschäftigt. Individuelle Archivierungsstrategien, eingeschränkte technische und organisatorische Kompetenzen und emotionale Entscheidungen verhindern automatisierte Archivierungshandlungen, wie sie für die institutionelle Langzeitarchivierung definiert sind. Diese Arbeit untersucht, welche Konzepte der institutionellen Langzeitarchivierung in skaliert Form auch durch private Anwender adaptiert werden können und wie diese in Deutschland durch Bibliotheken in ihrer Funktion als Experten für die institutionelle Langzeitarchivierung, aber auch als direkte Schnittstelle zum archivierenden Nutzer vermittelt werden können. Vorgeschlagen wird die Erstellung allgemeiner Informationen unter Beteiligung von nestor sowie die Vermittlung innerhalb von Veranstaltungen zur Informationskompetenz an wissenschaftlichen Bibliotheken und die praktische Umsetzung durch den Aufbau gemeinschaftlicher Repositorien an öffentlichen Bibliotheken.

ABSTRACT (ENGLISCH)

Personal digital archiving is an underresearched area of digital preservation that deals with the preservation of personal digital objects through individuals. Varying archiving practices, limited technical and organizational competencies as well as affective actions prevent the establishment of formal archiving standards. The thesis explores how digital preservation practices can be scaled down to benefit individuals' archiving needs and the role of German libraries as operating instances in digital preservation and information literacy. The distribution of basic archiving information for the general public through the nationwide nestor preservation cooperative is recommended as well as teaching personal digital archiving as a part of information literacy in higher education and the establishment of community archives through public libraries.

Schlagwörter: Personal Digital Archiving, Langzeitarchivierung, Informationskompetenz, öffentliche Bibliotheken, Personal Information Management, Community Archiving, digital preservation, information literacy, public libraries

INHALTSVERZEICHNIS

1	EINLEITUNG	1
2	PERSONAL DIGITAL ARCHIVING	3
2.1	Begriffsdefinition	3
2.2	Problemstellung	3
2.3	Wissenschaftliche Diskussion	5
2.4	Archivierungsverhalten	7
2.4.1	Wohlwollende Vernachlässigung bei unbegrenztem Speicherplatz	8
2.4.2	Speicherung außerhalb der lokalen Festplatte	9
2.4.3	Datenmigration	10
2.4.4	Bewertung von digitalen Objekten	12
2.4.5	Zusammenfassung	14
2.5	Konzepte und Modelle	14
2.6	Technische Unterstützung	16
2.6.1	Backup-Lösungen	16
2.6.2	MyLifeBits	16
2.6.3	HOPPLA	17
2.6.4	Zusammenfassung	18
2.7	Informations- und Schulungsangebote	18
2.8	Zusammenfassung	19
3	INSTITUTIONELLE LANGZEITARCHIVIERUNG	20
3.1	Begriffsdefinition	20
3.2	Problemstellung	20
3.3	Wissenschaftliche Diskussion	20
3.4	Archivierungsverhalten	22
3.4.1	Preservation Policy	22
3.4.2	Archivierungsstrategien	23
3.4.3	Verteilte Speicherung	25
3.5	Konzepte und Modelle	25
3.5.1	OAIS-Referenzmodell	26
3.5.2	Designated Community	27
3.5.3	Significant Properties	28
3.6	Technische Unterstützung	29
3.6.1	JHOVE	30
3.6.2	PLATO	31
3.6.3	Rosetta	32
3.7	Informations- und Schulungsangebote	32
3.8	Zusammenfassung	33

4	VERGLEICHENDE ANALYSE VON PERSONAL DIGITAL ARCHIVING UND INSTITUTIONELLER LANGZEIT-ARCHIVIERUNG	34
4.1	Problemstellung	35
4.2	Wissenschaftliche Diskussion	35
4.3	Archivierungsverhalten	36
4.3.1	Archivierungsvorschriften	36
4.3.2	Wert	36
4.3.3	Speicherorte und Speicherverhalten	38
4.3.4	Speichern/Langzeitarchivieren	39
4.3.5	Daten- und Formatmigration	39
4.4	Konzepte und Modelle	40
4.5	Technische Unterstützung	40
4.6	Informations- und Schulungsangebote	41
4.7	Fazit der Analyse	42
5	VORSCHLÄGE FÜR DIE ZUKÜNFTIGE VERMITTLUNG VON PERSONAL DIGITAL ARCHIVING IN DEUTSCHLAND	43
5.1	Personal-Digital-Archiving-Hinweise für die breite Öffentlichkeit	43
5.1.1	InterPARES2 Creator Guidelines: Making and maintaining digital materials: Guidelines for individuals	44
5.1.2	PARADIGM Guidelines for creators of personal archives	45
5.1.3	NDIIPP Preserving your digital memories	45
5.1.4	Fazit	47
5.2	Personal Digital Archiving als Teil der Informationskompetenzvermittlung an wissenschaftlichen Bibliotheken	48
5.3	Personal Digital Archiving an öffentlichen Bibliotheken	49
6	FAZIT	52
7	QUELLEN- UND LITERATURVERZEICHNIS	54
	ANHANG	

ABBILDUNGSVERZEICHNIS

Abbildung 1: Funktionales OAIS-Referenzmodell	26
Abbildung 2: Installationsverzeichnis und grafische Oberfläche von JHOVE	30
Abbildung 3: Startseite von Plato 4.4	31

TABELLENVERZEICHNIS

Tabelle 1:	Fallbeispiel Gewichtung von Significant Properties	29
Tabelle 2:	Gegenüberstellung der unterschiedlichen Ansätze	34
Tabelle 3:	Berücksichtigung der primären Personal-Digital-Archiving-Maßnahmen in den einzelnen Anleitungen	47

ABKÜRZUNGSVERZEICHNIS

ACRL	Association of College and Research Libraries
EU	Europäische Union
HOPPLA	Home and Office Painless Persistent Long-term Archiving
InterPARES	International Research on Permanent Authentic Records in Electronic Systems
ISO	International Organization for Standardization
JHOVE	JSTOR/Harvard Object Validation Environment
JSTOR	Journal STORage
LoC	Library of Congress
LOCKSS	Lots Of Copies Keep Stuff Safe
NAS	Network Attached Storage
NDIIPP	National Digital Information Infrastructure and Preservation Program
nestor	Network of Expertise in long-term Storage and availability of digital Resources in Germany
NSA	National Security Agency
OAIS	Open Archival Information System
PARADIGM	Personal Archives Accessible in Digital Media
PIM	Personal Information Management
PLANETS	Preservation and Long-term Access through Networked Services
PLATO	Preservation Planning Tool
RAID	Redundant Array of Independent Disks

SCAPE	Scalable Preservation Environments
TU	Technische Universität
UK	United Kingdom
USA	United States of America
4C	Collaboration to Clarify the Costs of Curation

1 EINLEITUNG

Die Lebenswelt der Menschen in Deutschland und anderen westlichen Staaten ist mittlerweile durch zunehmende Digitalisierung ehemals analoger Vorgänge geprägt. Briefe werden durch E-Mails abgelöst, Fotos nur noch digital erstellt und Anträge können auf den Webseiten der kommunalen Verwaltungen online ausgefüllt werden, um nur einige Beispiele zu nennen. Jeder Bürger¹ verfügt mittlerweile über eine Menge von Daten, die ihm nur noch in digitaler Form zugehen bzw. die er selber digital erstellt. Dabei handelt es sich sowohl um persönliche Daten wie Fotos, Musik, Videos, private E-Mails oder Kurznachrichten, aber auch um offizielle Dokumente wie Verträge, Steuererklärungen, Reiseunterlagen, Versicherungsscheine etc., die mittlerweile auch nur noch selten per Briefpost zugestellt werden, da sich auch hier z. B. der E-Mail-Versand als PDF-Attachement durchsetzt.

Die digitale Übermittlung und Speicherung erlaubt in der Regel die Beschleunigung von Vorgängen und einen schnelleren Zugriff auf die benötigten Daten. In den Hintergrund tritt hierbei die Fragilität dieser Daten. Digitale Daten sind vielen Risiken ausgesetzt. Dazu gehören Verlust in Folge von Festplattencrashes, Abhängigkeit von bestimmten Softwareprodukten zur Nutzung der Datei, Formatänderungen oder der sogenannte Bit Rot, die zeitlich bedingte Zerstörung der Datei. Aber auch die Menge der digitalen Objekte, ihre verstreute Speicherung auf Festplatten und in der Cloud und fehlende Metadaten verhindern das Auffinden bestimmter Objekte, wodurch sie ebenfalls als verloren angesehen werden müssen.

Die meisten Menschen sind sich aber dieses Problems der Fragilität ihrer Daten gar nicht bewusst; im Gegenteil meint man mit der Digitalisierung analoger Objekte wie z. B. Fotos ihre Lebenszeit zu verlängern. Es besteht kein Bewusstsein dafür, dass man digitale Daten managen muss, um sie zu erhalten. Andere kennen zwar das Problem, scheuen aber die Beschäftigung mit den ihnen fremden technischen Gegebenheiten und Maßnahmen zur Erhaltung ihrer Dateien bzw. vernachlässigen diese aus Zeitgründen. Dadurch gehen viele Daten verloren, die für den Einzelnen persönlichen Wert haben bzw. im Falle von Verträgen etc. auch eine rechtliche Bewandnis haben.

Private Daten machen einen Großteil der weltweiten Datenmenge aus. Organisationen wie Bibliotheken wenden auf ihre digitalen Bestände umfangreiche Langzeitarchivierungsmaßnahmen an und treiben deren Entwicklung weiter voran, der breiten Öffentlichkeit und dem persönlichen Archivierungsbedarf des Einzelnen kommen diese Erkenntnisse bisher aber kaum zu Gute. Während öffentliche Daten mit hohem Aufwand erhalten werden, verlieren die Bürger viele ihrer für sie wertvollen Daten.

Die wissenschaftliche Beschäftigung mit der Verwaltung und Archivierung von persönlichen digitalen Daten durch Privatpersonen findet unter dem Begriff Personal Digital Archiving statt. Studien zum Archivierungsverhalten zeigen eklatante Unterschiede zu Konzepten der Langzeitarchivierungsforschung. Durch Information und Schulung könnten diese Archivierungshandlungen verbessert werden und private Daten länger erhalten bleiben.

Bibliotheken als Betreiber institutioneller Langzeitarchivierung und als Informations- und Schulungsanbieter im Bereich der digitalen Informationsrecherche und Informationsverwaltung mit direktem Kontakt zum privaten Anwender scheinen prädestiniert für die Erstellung von Informationsangeboten zur Archivierung persönlicher Daten, da sie sowohl über das fachliche Know-how zu Langzeitarchivierungsmethoden als auch über etablierte Schulungs- und Informationsangebote verfügen. Außerdem können durch Bibliotheken sowohl Nutzer wissenschaftlicher Bibliotheken, als auch die Nutzer öf-

¹ Das generische Maskulinum impliziert hier und nachfolgend die Gleichwertigkeit männlicher und weiblicher Akteure.

fentlicher Bibliotheken erreicht werden. Das vorhandene Wissen zur institutionellen Langzeitarchivierung sollte unter Beachtung der Forschungsergebnisse zu Personal Digital Archiving entsprechend aufbereitet werden, damit es Einzelpersonen vermittelt werden kann.

Die vorliegende Arbeit untersucht, welche Besonderheiten der Archivierung persönlicher Daten durch Privatanwender beachtet werden müssen, welche Erkenntnisse aus dem Bereich der institutionellen Langzeitarchivierung genutzt werden können und in welcher Form entsprechende Informations- und Schulungsangebote konzipiert werden können.

In Kapitel 2 wird einleitend der Begriff Personal Digital Archiving, die besondere Problemstellung und die bisherige wissenschaftliche Diskussion erläutert. Danach folgt eine Darstellung des bisher in Studien untersuchten Archivierungsverhaltens von Privatpersonen und der daraus folgenden Implikationen für theoretische Modelle, technische Unterstützung durch Softwarelösungen und Informations- und Schulungsangebote.

Kapitel 3 stellt analog zum Vorgehen in Kapitel 2 Problemstellung und wissenschaftliche Diskussion der institutionellen Langzeitarchivierung dar, gefolgt von entsprechenden Erläuterungen zu Archivierungsverhalten, Theorien und Modellen sowie technischen Lösungen und Schulungsangeboten.

In Kapitel 4 erfolgt dann die vergleichende Analyse der beiden Forschungsdisziplinen, mit dem Ziel, Ansätze für die Verbesserung der derzeitigen Personal Digital Archiving Praktiken zu evaluieren.

In Kapitel 5 werden schließlich konkrete Ansätze für die Erstellung von Informationsangeboten durch Bibliotheken für die breite Öffentlichkeit, Nutzer wissenschaftlicher Bibliotheken und Nutzer öffentlicher Bibliotheken skizziert.

Betont werden sollte an dieser Stelle, dass lediglich untersucht wird, welche Angebote für Menschen geschaffen werden können, die bereits ein Interesse an der Langzeitverfügbarkeit ihrer Daten entwickelt haben – offen bleiben wird die Frage, wie man Privatmenschen trotz des zu erwartenden Aufwands motivieren kann, tatsächlich Personal Digital Archiving zu betreiben.

Die dieser Arbeit zugrundeliegende Perspektive ist die des privaten Anwenders. Es soll nicht die Perspektive von Bibliotheken eingenommen werden, die private Nachlässe in digitaler Form nach dem Tod der entsprechenden Person in ihren Bestand übernehmen und somit ein Interesse an der Übergabe der Daten in langzeitarchivierungstauglicher Form haben, sondern Bibliotheken sollen den privaten Anwender objektiv unterstützen, so dass er selber in die Lage versetzt wird, seine Daten gemäß den ihm zur Verfügung stehenden Möglichkeiten und zu dem ihm sinnvoll erscheinenden Zweck zu bewahren.

Juristische, insbesondere urheberrechtliche Fragen hinsichtlich Personal Digital Archiving werden nicht näher erläutert, da hierbei ein komplexes Rechtsgebiet Anwendung findet. Daher liegt der Fokus der Arbeit insbesondere auf der Bewahrung selbstgestellter digitaler Objekte, für die der Nutzer die entsprechenden Urheberrechte besitzt; die Bewahrung von digitalen Objekten, an denen der Nutzer ausschließlich ein Nutzungsrecht erworben hat, ist für diese Arbeit sekundär.

2 PERSONAL DIGITAL ARCHIVING

Die Antwort auf die Frage, was Personal Digital Archiving ist, könnte ganz knapp lauten: Personal Digital Archiving bezeichnet das Archivieren digitaler Objekte durch Privatpersonen. In der Vorbereitungsphase dieser Arbeit wurde diese Frage jedenfalls öfter in Flurgesprächen mit Kollegen gestellt und mit obigem Satz beantwortet. Die Reaktionen auf diese knappe Erklärung fielen dabei ganz unterschiedlich aus: Die einen verstanden Personal Digital Archiving als reines Datensicherungsthema und klopfen sich selber auf Grund ihrer vorbildlichen Maßnahmen auf die Schulter, die anderen zweifelten, was denn daran nun wichtig genug sein sollte, um darüber eine Masterarbeit zu schreiben. Beide Reaktionen zusammengenommen zeigen, dass zwar schon eine Ahnung vorhanden ist, dass die Archivierung persönlicher bzw. privater digitaler Objekte notwendig ist, ein Problembewusstsein aber derzeit fehlt.

Auch bei der Erwähnung des Themas dieser Arbeit gegenüber im Bibliotheksbereich tätigen Gesprächspartnern reichten die Vorstellungen von dem was Personal Digital Archiving ist von der Archivierung von Personalakten über Selbstarchivierung bis hin zu e-Portfolios. Dies zeigt, dass der Begriff und seine Definition auch in Personenkreisen wenig bis gar nicht bekannt ist, die sich von Berufs wegen eigentlich mit der Erhaltung und Archivierung auch von digitalen Objekten beschäftigen.

Daher soll im Folgenden der Begriff Personal Digital Archiving und das dahinterstehende Problem erläutert werden.

2.1 Begriffsdefinition

Personal Digital Archiving ist aus dem im englischen Sprachraum verwendeten Begriff Personal Archiving hervorgegangen, der die durch Privatpersonen vorgenommene Archivierung von für sie wichtigen analogen Objekten wie Fotos, Unterlagen etc. bezeichnet. Personal Digital Archiving überträgt diesen Begriff nun auf die Archivierung ausschließlich digital vorliegender Objekte, seien diese born digital, also direkt digital erzeugt oder digitalisiert, wie z. B. Scans von analogen Fotos. Allerdings ist Personal Digital Archiving (noch) kein feststehender Begriff; die Library of Congress z. B. verwendet auf ihren Informationsseiten² den Begriff Personal Digital Archiving alternierend mit Personal Archiving, obwohl ausdrücklich auf die Archivierung digitaler Objekte eingegangen wird. Daher wird innerhalb dieser Arbeit der Begriff Personal Digital Archiving verwendet; auf die in der englischsprachigen Forschung genutzte Abkürzung PDA wird aufgrund der etablierten Verwendung derselben für das Konzept der Patron-Driven Acquisition verzichtet.

2.2 Problemstellung

Von der bis zum Jahr 2010 erzeugten digitalen Datenmenge wurden 70 % durch Privatpersonen erstellt.³ Das bedeutet, dass der größte Teil der weltweiten Datenmenge nicht durch Firmen und Institutionen erzeugt wird, die über entsprechende Archivierungs- und Datensicherungsstrategien verfügen, sondern durch Privatleute am heimischen PC, die oftmals keine konkrete Datensicherung und Archivierung betreiben.

Die meisten Datenverluste sind aber nicht durch Festplattencrash und andere Hardwarebeschäden bedingt, denen man mit entsprechenden Backupstrategien begegnen kann: Laut der Digital-Lives-Studie resultieren etwa 70 % des Datenverlusts aus der Unauf-

² Vgl. Library of Congress o.J.a

³ Vgl. Marshall 2013, S. 94

findbarkeit der gesuchten Datei.⁴ Das bedeutet, dass die gesuchten Daten in der Regel sogar noch physisch vorhanden sind, es aber nicht möglich ist, diese zu finden, z. B. weil sie in der Menge der lokal gespeicherten Daten, die oft ohne strukturierte Ablage und Dateinamenvergabe auskommen muss, nur durch Autopsie aller in Frage kommenden Dateien aufgespürt werden können, indem sich der Suchende mühsam durch sämtliche Ordner und Dateien klickt. Aber auch die zunehmende Speicherung von Dateien außerhalb der lokalen Festplatte z. B. mittels Cloudspeicherdiensten wie Dropbox oder die direkte Speicherung von Daten in sozialen Netzwerken wie Facebook oder auf Plattformen wie Flickr und YouTube ohne den Weg über die Kopie auf der lokalen Festplatte tragen zur Unauffindbarkeit und damit zum effektiven Verlust persönlicher Daten bei.

Das eigentliche Problem hinter Personal Digital Archiving ist also nicht unbedingt ein technisches, sondern eher ein Problem der Organisation und Verwaltung persönlicher Daten durch den Nutzer. In diesem Zusammenhang trifft man in der Literatur oft auf den Begriff der „digital curation“ bzw. „personal curation“:

Digital curation involves maintaining, preserving and adding value to digital research data throughout its lifecycle.⁵

Diese Definition bezieht sich insbesondere auf das Management von Forschungsdaten, die einen bestimmten Lebenszyklus durchlaufen: sie werden erhoben, genutzt, archiviert und dann eventuell von anderen Forschern nachgenutzt. Aber auch private Daten werden erstellt, genutzt, archiviert und eventuell später wieder benötigt, wieso der abgeleitete Ausdruck „personal curation“ durchaus zutreffend ist.

Hinderlich für die Wahrnehmung von Personal Digital Archiving als Problem ist die allgemein propagierte Annahme, dass analoge Fotos, Filme etc. schon durch das Digitalisieren an sich „gerettet“ sind und erhalten bleiben. Ein anschauliches Beispiel hierfür ist z. B. das vom WDR betriebene „Archiv des analogen Alltags“, *Digit*.⁶ Auf dieser Plattform können Nutzer ihre Fotos und selbstgedrehten Filme hochladen, um diese mit anderen zu teilen und für die Nachwelt zu „retten“, wie es in der Beschreibung heißt:

Digit ist das digitale Archiv des analogen Alltags. Auf dieser Plattform wollen wir Fotos und Filme aus Zeiten sammeln, als Bilder noch entwickelt werden mussten und Filmen noch ein teures Hobby war. Hier kann jeder beim Blättern in Erinnerungen schwelgen oder über die verrückten, vergangenen Jahrzehnte den Kopf schütteln. Wir möchten jeden ermutigen, im eigenen Keller oder auf dem Dachboden in Kisten zu wühlen, Fotos und Filme zu retten und hier mit allen anderen zu teilen – werde ein Digger!⁷

Weiter heißt es:

Digger sind Menschen wie du und ich. Sie machen sich auf die Suche nach prädigitalen Fundstücken, um sie vor dem Verfall zu retten. Und um sie mit möglichst vielen ihrer Mitmenschen zu teilen [to dig, Verb, engl. – graben, ausgraben, buddeln]⁸

⁴ Vgl. John et al. 2010, S. 23

⁵ Digital Curation Centre

⁶ WDR 2014c

⁷ WDR 2014a

⁸ WDR 2014a

Man mag dem WDR sein ehrenvolles Engagement für die Rettung von „prädigitalen Fundstücken“ abnehmen oder nicht – die hochgeladenen Bilder und Filme werden vom WDR z. B. als historisches Material für das tägliche Nachrichtenmagazin *Aktuelle Stunde* verwendet, die entsprechenden Nutzungsrechte räumt der Nutzer mit dem Upload der Bilder bzw. Filme ein⁹ – in jedem Fall wird auch durch den oben wiedergegebenen Wortlaut der Gemeinplatz bedient, dass durch die Digitalisierung allein ein analoges Objekt bereits gerettet ist. Dabei wird erst gar nicht die Frage gestellt, ob das analoge Objekt eventuell bei entsprechender Lagerung und – im Anwendungsfall Film – der Erhaltung der benötigten Abspieltechnik sogar länger erhalten bleiben werden könnte als eine digitale Datei, die durch Hardwareausfall verloren geht oder in der Menge der digital gespeicherten Daten auf den immer größer werdenden Speichermedien verschwindet. Auf der Plattform selber werden weder Hinweise zur Verfügbarkeitsdauer des hochgeladenen Materials oder zur Langzeitarchivierung durch den WDR gegeben, noch wird der Nutzer darauf hingewiesen, wie er selber das hochgeladene bzw. durch den WDR digitalisierte Material archivieren sollte. Die Erhaltung und „Rettung“ hört mit der Digitalisierung bzw. mit dem Upload des Materials auf die Plattform auf. Dabei müssen auch Digitalisate verwaltet werden, damit sie möglichst lange erhalten bleiben und nicht mit dem nächsten generellen Formatwechsel im Bereich Foto/Video eine erneute Rettungsaktion nötig wird.

Dass dieses Problem der Verwaltung und Pflege von digitalen Objekten sowohl von normalen Anwendern als auch von Anwendern aus dem Bibliotheksbereich nicht direkt erfasst wird, hat unterschiedliche Ursachen. Zum einen wird den meisten Nutzern das Archivierungsproblem erst bewusst werden, wenn tatsächlich ein Datenverlust auftritt, es also in der Regel bereits zu spät für entsprechende vorbeugende Maßnahmen ist. Zum anderen ist das Speichern digitaler Objekte etwas so Alltägliches und gleichzeitig Mittel zum Zweck, dass eine wissenschaftliche Beschäftigung hiermit erst gar nicht in Erwägung gezogen wird. Dem widerspricht allerdings die vorwiegend in den USA geführte wissenschaftliche Diskussion der letzten Jahre zum Thema Personal Digital Archiving, die im folgenden Abschnitt vorgestellt wird.

2.3 Wissenschaftliche Diskussion

In der Fachliteratur wird der Begriff Personal Digital Archiving neben anderen Bezeichnungen wie Personal Collection, Personal Curation, Personal Record Management etc. verwendet. Daher ist es schwierig zu bestimmen, zu welchem Zeitpunkt die Archivierung digitaler Objekte von und durch Einzelpersonen in die wissenschaftliche Diskussion Einzug gefunden hat.

In seinem Artikel im *dlib Magazine* im Juni 2005 begründet der Langzeitarchivierungsberater Neil Beagrie seine Motivation für die wissenschaftliche Beschäftigung mit der Archivierung durch Privatpersonen mit zwei Trends, die sich bis heute fortsetzen: Zum einen die auch als Mooresches Gesetz beschriebene Verdopplung der Rechenleistung alle 18 Monate, die sich auch in Verfügbarkeit und Preis von Speichermedien niederschlägt, zum anderen der Trend zur Erstellung von Inhalten durch Nutzer, also das, was gemeinhin als Web 2.0 bezeichnet wird und was sich nach 2005 durch Plattformen wie Facebook und Twitter noch erweitern sollte.¹⁰

Diese beiden Trends hatten sich nach Beagries Ansicht im Jahr 2005 so weit entwickelt, dass sowohl Speichermedien mit großer Kapazität als auch die Erstellung von Inhalten durch den bisher rein rezipierenden Nutzer attraktiv, erschwinglich und nutzbar

⁹ Vgl. WDR 2014b

¹⁰ Vgl. Beagrie 2005

geworden waren, weshalb zu diesem Zeitpunkt das Nachdenken über Archivierung durch Endnutzer für ihn in den Blickpunkt rückte.¹¹

Tatsächlich sind seit 2005 durchgehend Aufsätze, Beiträge und Kommentare recherchierbar, die sich mit dem Thema der Archivierung durch Endnutzer beschäftigen. Regelmäßige Veröffentlichungen von Studien zum Archivierungsverhalten liefert z. B. Cathy Marshall von Microsoft Research; einige dieser Erkenntnisse zum Nutzerverhalten werden im folgenden Kapitel genauer erläutert werden. Cathy Marshalls Arbeiten werden dem Forschungsbereich Personal Information Management zugeordnet, der sich mit der Organisation sowohl analoger als auch digitaler Information im privaten Umfeld beschäftigt. Die klassische archivarische Sicht auf das Thema Personal Digital Archiving wird z. B. durch Amber Cushing vertreten, die die Ergebnisse aus Cathy Marshalls Studien archivarischen Ergebnissen und Methoden gegenüberstellt.¹² Stimulus für diese explizit archivarische Beschäftigung mit dem Thema Personal Digital Archiving war die Aussage Marshalls, dass Methoden und Theorien der institutionellen Archivierung nur auf Teilaspekte von Personal Digital Archiving angewandt werden können.¹³ Auch Peter Williams und Kollegen greifen diese Aussage auf und untersuchen Personal Digital Archiving aus archivarischer Sicht.¹⁴ Dabei bestätigen sowohl Cushing als auch Williams et al. die Forschungsergebnisse Marshalls und plädieren für eine weitere archivarische Behandlung des Themas. Allerdings bleiben sie die Antwort schuldig, welchen Effekt die archivarische Auseinandersetzung mit dem Thema denn genau haben soll; konkrete, praktische Erkenntnisse werden in beiden Veröffentlichungen nicht genannt. Immerhin regt Cushing an:

Future research on personal digital archiving may benefit from collaborations between members of the PIM community and members of the archival community.¹⁵

Aus bibliothekarischer Sicht wird das Thema von Andrea J. Copeland untersucht, die in einer Studie¹⁶ das Archivierungsverhalten von Nutzern öffentlicher Bibliotheken untersuchte; auch diese Ergebnisse werden im folgenden Kapitel Erwähnung finden.

Recherchiert werden konnten außerdem einzelne größere Forschungsprojekte zum Thema Personal Archiving im Zeitraum seit 2005, wie z. B. das bereits zitierte Digital-Lives-Projekt der British Library. Dieses Projekt beschäftigt sich mit Personal Digital Archiving aus der Perspektive einer späteren Übernahme persönlicher digitaler Kollektionen in institutionelle Repositorien:

“Digital Lives” is a research project focusing on personal digital collections and their relationship with research repositories.¹⁷

Dazu wurden sowohl die Archivierungsmethoden von Wissenschaftlern als auch von Menschen außerhalb der universitären Forschung untersucht.¹⁸ Leider ist das Projekt nur in einigen wenigen Veröffentlichungen dokumentiert. Die zugehörige Website¹⁹ ist nicht mehr auffindbar bzw. nur noch in archivierter Form vom 23. April 2012 im UK

¹¹ Vgl. Beagrie 2005

¹² Cushing 2010

¹³ Vgl. Cushing 2010, S. 302; Marshall 2007, S. 4

¹⁴ Williams et al. 2009

¹⁵ Cushing 2010, S. 310

¹⁶ Copeland 2011

¹⁷ Williams et al. 2008

¹⁸ Williams et al. 2008

¹⁹ Siehe The British Library 2014

WebArchive verfügbar²⁰, der Abschlussbericht ist nur in verschiedenen Beta-Versionen²¹ online auffindbar. Somit ist eine Nachnutzbarkeit der Ergebnisse nur erschwert möglich.

Ein weiteres im UK beheimatetes Projekt war PARADIGM²², das sich von 2005 bis 2007 ebenfalls mit der Übernahme privater digitaler Kollektionen in institutionelle Repositorien beschäftigte und dies an Hand der digitalen Kollektionen zweier Politiker untersuchte. Im Gegensatz zum Digital-Lives-Projekt ist hier eine ausführliche Dokumentation in Form eines Workbooks vorhanden, das Arbeitsabläufe detailliert beschreibt und die praktische Anwendung in den Vordergrund stellt.²³

Die Library of Congress hält seit 2010 auf ihrer Website²⁴ entsprechende Informationen zu Personal Digital Archiving bereit und verweist auch auf die oben genannten Projekte. Außerdem unterstützt sie die seit 2010 jährlich stattfindende Konferenz zum Thema Personal Digital Archiving.²⁵ Ziel dieser Konferenz ist es auch, Wissenschaftler mit unterschiedlichem Forschungsbackground zusammenzubringen, um wie auch von Amber Cushing angeregt, den interdisziplinären Austausch zu fördern.²⁶

Auch aus der technischen Perspektive wurde das Thema in den letzten Jahren erkundet, z. B. durch Forscher der TU Wien im Rahmen des EU-Langzeitarchivierungsprojekts PLANETS²⁷, in dem für verschiedene Langzeitarchivierungsaufgaben entsprechende Software-Lösungen bereitgestellt wurden. Ihre Lösung HOPPLA²⁸ richtet sich an kleine und mittlere Unternehmen sowie Privatanwender und wird in Kapitel 2.6.3 vorgestellt.

Die wissenschaftliche Diskussion deckt also viele Perspektiven und Interessenlagen ab; trotzdem ist das Thema Personal Digital Archiving in Deutschland nahezu unbekannt. Suchanfragen mit dem Begriff Personal Digital Archiving in Google ergeben keine relevanten deutschsprachigen Treffer. Nicht besser sieht es bei der Recherche in der virtuellen Fachbibliothek der Buch-, Bibliotheks- und Informationswissenschaft, b2i, aus, auch hier sind die Treffer in der Regel englischsprachig. Die Nutzung möglicher deutscher Suchterme wie z. B. „persönliche Archivierung“ liefert ebenfalls keine Treffer zum Thema. Alle gefundenen Resultate beschäftigen sich mit institutioneller Archivierung z. B. durch Bibliotheken oder Archive bzw. mit eher kaufmännischen Fragen wie Compliance und Gewährleistung als Grundlage für die Archivierung von digitalen Objekten.

Auch bei der Deutschen Nationalbibliothek scheint das Thema Personal Digital Archiving nicht im Fokus zu stehen: Im Gegensatz zur Library of Congress bietet die Deutsche Nationalbibliothek auf ihrer Website keine Archivierungshinweise für die private Archivierung digitaler Daten an.

2.4 Archivierungsverhalten

Gegenstand der Untersuchungen von Marshall, Williams et al. und Copeland ist das Archivierungsverhalten einzelner Menschen hinsichtlich digitaler Objekte. Die generellen Fragen hierbei sind, inwiefern sich das Archivierungsverhalten einzelner Menschen

²⁰ Siehe The British Library o.J.

²¹ John et al. 2010 scheint die aktuellste Version (Beta 0.2) zu sein.

²² Personal Archives Accessible in Digital Media; siehe PARADIGM 2005a

²³ PARADIGM 2005b

²⁴ Siehe Library of Congress o.J.a

²⁵ Indiana State University

²⁶ Vgl. Kim 2013, S. 174

²⁷ Siehe PLANETS 2007

²⁸ HOPPLA ist das Akronym von Home Office Painless Persistent Long-Term Archiving

unterscheidet, wovon dies abhängig ist (technische/konzeptuelle/emotionale Faktoren) und ob und wie sich hieraus generelle Verhaltensweisen ableiten lassen, die als Basis für eine Verbesserung der Archivierung sowohl durch bestimmte Methoden als auch durch unterstützende Softwarelösungen genutzt werden können.

2.4.1 Wohlwollende Vernachlässigung bei unbegrenztem Speicherplatz

Ein in der Literatur häufig bemühtes Bild für das digitale Archivierungsverhalten von privaten Anwendern ist das des digitalen Schuhkartons.

Beinahe jeder Mensch sammelt und archiviert bestimmte Dinge, die ihm wichtig erscheinen, in der Regel aus emotionalen Gründen. Dies kann z. B. die Schuhschachtel mit Familienfotos sein, Bücher, Tonträger oder alte Konzertkarten als Erinnerung an einen schönen Abend. In der Regel überleben diese Memorabilien Umzüge und andere örtliche Veränderungen, fristen ihr Dasein in dunklen Kellern oder auf feuchten Dachböden. Der Akt des Wegwerfens scheitert oft: Die haptische Begegnung mit dem Sammelobjekt lässt in der Regel Erinnerungen wieder aufkommen und so überstehen die meisten dieser Dinge Frühjahrsputz und Aufräumaktionen unversehrt. Oder der Aufräumende bemerkt, dass für das Aussortieren der Sammlung entsprechend Zeit aufgewendet werden müsste und verschiebt daher das Aussortieren auf einen künftigen Zeitpunkt.

Ganz ähnlich verhält es sich mit digitalen Objekte: Auch hier ist das Archivierungsverhalten durch „benign neglect“²⁹, also wohlwollende Vernachlässigung gekennzeichnet, wie es Cathy Marshall nennt:

It is not unusual for consumers to write the most valuable of their files to external media and to put that media – a succession of floppy disks, zip drives, external hard drives, CDs, DVDs, and other storage solutions of the moment – aside in a safe place, possibly at a physical remove (such as a safety deposit box). Of the ways people archive now, this one most closely matches the kind of benign neglect with which we archive physical materials like photos and important paper documents. We set aside the things that matter in a special place, and hope they will be there when the time comes to retrieve them.³⁰

Die externe Festplatte wird also als digitaler Schuhkarton genutzt, der größtenteils ungeordnet digitale Artefakte wie z. B. Familienfotos speichert, die dann vielleicht irgendwann einmal sortiert und verwendet werden.

Nun besteht zwischen einem handelsüblichen Schuhkarton und einer USB-Festplatte jedoch ein großer Unterschied hinsichtlich des Fassungsvermögens: Während der Schuhkarton begrenzten Platz für die Aufnahme von Fotos bietet, wird das Speichervolumen von externen Festplatten immer größer und günstiger. Aktuell kosten drei Terabyte Speicherplatz circa 100 Euro. Damit steht den Verbrauchern zur Speicherung persönlicher Daten scheinbar unbegrenzt Speicherplatz zur Verfügung. Jedoch wirkt sich die Verfügbarkeit von Speicher natürlich auf das Archivierungsverhalten aus; wie bereits weiter oben zitiert, sieht Neal Beagrie die Verfügbarkeit von Speicherplatz zu einem angemessenen Preis überhaupt erst als Auslöser persönlichen Archivierens. Dadurch, dass nun scheinbar unbegrenzt Speicher verfügbar ist, wird nur in seltenen Fällen darüber nachgedacht, ob eine Datei wirklich speicherungswürdig ist – in der Regel wird alles gespeichert.

²⁹ Marshall 2008a

³⁰ Marshall 2008a

Diese Tendenz besteht nicht nur im privaten Bereich: In den Anfangszeiten der Software-Entwicklung durften Programme nicht größer als 256 Kilobyte sein, schlicht und einfach weil nicht mehr Hauptspeicher zur Verfügung stand. Heutzutage käme sicherlich kein Programm mehr mit derart geringem Speicher aus und es bestünde auch gar kein Grund, Programme derart speichereffizient anzulegen, da die gängige Endanwendungshardware über entsprechende Kapazitäten verfügt.

Analog hierzu werden auch Nutzer, die bisher nur die durchschnittliche interne Speicherkapazität ihres Computers genutzt haben, nach der Anschaffung einer externen Festplatte mit der zehnfachen Speicherkapazität ihr Archivierungsverhalten ändern, weil eben nicht mehr mit dem vorhandenen Speicherplatz gewirtschaftet werden muss. Für den Prozess des Wiederauffindens verlassen sich die Nutzer auf ihr Erinnerungsvermögen, Browsing und Desktopsuchmaschinen, wie die Studie von Copeland belegt:

Browsing to refind personal content was the norm (n=21), likely owing to the limitations imposed by digital devices on metadata creation and search functionality.³¹

Das Archivierungsverhalten ist also durch das Abwägen von Kosten und Nutzen bestimmt. Beagrie drückt dies so aus:

The combination of cheap digital storage and very sophisticated retrieval tools is shifting the balance of costs: digitally it is becoming cheaper to collect and more expensive to select, and cheaper to search than to organize.³²

Ob diese Rechnung zu dem Zeitpunkt aufgeht, an dem man eine Datei wirklich wiederfinden möchte, darf bezweifelt werden. Desktopsuchmaschinen weisen bei der Auffindung von Dateien große Unzulänglichkeiten auf. Und wie sollen Dateien wiedergefunden werden, von denen der Suchende gar nicht mehr weiß, dass er sie überhaupt hat?

2.4.2 Speicherung außerhalb der lokalen Festplatte

Hinzu kommt, dass nicht nur die Hardware leistungsfähiger wird hinsichtlich der Speicherkapazitäten, sondern mittlerweile Online-Services ermöglichen, dass Dateien auch außerhalb des eigenen Computers gespeichert werden. Auf diese verteilte Speicherung, z. B. von Fotos auf Flickr oder auf anderen Social-Media-Plattformen zum Zwecke des Teilens mit anderen Nutzern, aber auch zum Zweck der dauerhaften Archivierung geht Cathy Marshall schon 2008 ein.³³ Inzwischen dürfte sich diese Praxis der dauerhaften Speicherung im Internet durch zwei neuere Plattformen mit noch größerem Einfluss auf die tägliche Speicherung von Dateien noch verschärft haben: Facebook und Dropbox.

Facebook erlaubt das Teilen von Fotos und Videos mit Freunden und unterstützt insbesondere das direkte Hochladen dieser Dateien vom Smartphone. Ähnlich arbeitet Dropbox: Auch hier ist es möglich, mit dem Smartphone aufgenommene Fotos und andere Dateien direkt in Dropbox zu speichern. Dropbox ermöglicht aber noch mehr: An einem Arbeitsplatz gespeicherte Dateien werden automatisch im cloudbasierten Speicher aktualisiert und können für die Benutzung an anderen Arbeitsplätzen synchronisiert werden, so dass man von überall auf der Welt immer auf die aktuellste Version einer Datei zugreifen kann. Außerdem können Dateien für andere Nutzer freigegeben werden, so dass kollaboratives Arbeiten ermöglicht wird. Versionskonflikte werden

³¹ Copeland 2011, S. 1295

³² Beagrie 2005

³³ Siehe hierzu Marshall 2008a und b

dabei erkannt und durch automatische Speicherung unter einem angepassten Dateinamen umgangen.

Diese verteilte Speicherung und insbesondere die bei Dropbox integrierten Synchronisationsmöglichkeiten geben dem Nutzer eine bestimmte Sicherheit hinsichtlich der Speicherung seiner Dateien – Verlust muss hier nicht befürchtet werden.

Aber ist das auch Archivierung? Cathy Marshall bezeichnet die verteilte Speicherung von Dateien als ungenügend, weder werden Inhalte vor der Sicherung bewertet, noch ist gewährleistet, dass diese Dateien wirklich für einen langfristigen Zeitraum zur Verfügung stehen – Dienste werden eingestellt, Benachrichtigungen über die Einstellungen eines Dienstes werden an nicht mehr genutzte E-Mail-Adressen gesendet, Passwörter für bestimmte Dienste werden vergessen.³⁴ Damit sind auch diese verteilten Daten verloren, wenn keine weiteren Kopien zur Verfügung stehen. Und auch bei den weiteren Kopien ist nicht unbedingt gesagt, dass diese dem verlorenen Original entsprechen: Zum Beispiel werden Fotos für die Darstellung auf bestimmten Plattformen im Internet hinsichtlich Größe und Auflösung angepasst, die digitale Datei wird also verändert und erleidet eventuell Qualitätsverluste gegenüber dem Original.

Archivierung impliziert, dass Daten nicht ad-hoc genutzt werden, sondern dass sie an einem bestimmten, von der frequenten aktiven Nutzung ausgeschlossenen Ort bereitgehalten werden. Dadurch dass die Begriffe Speichern und Archivieren oft synonym genutzt werden, kommt es wahrscheinlich auch zu dem Missverständnis, dass reine Backups bereits als Archivierungslösung betrachtet werden, wie es auch Cathy Marshall in ihren Studien beobachtet. Dabei übersehen die meisten Nutzer, dass Backups, wenn sie überhaupt stetig durchgeführt werden, zwar im Hier und Jetzt eine gute Möglichkeit sind, sich vor Datenverlust zu schützen, dies aber eben keine Langzeitarchivierung ist – und die wenigsten überhaupt einmal einen Restore eines Systembackups aktiv durchgeführt haben.³⁵

Ein weiteres Problem der Online-Speicherung ist, dass vielen Menschen überhaupt nicht bewusst ist, dass diese Daten nicht lokal auf ihrem Computer gespeichert sind, sondern nur auf einem entfernten Server bereitgehalten werden und für den Zugang immer eine Internetverbindung benötigt wird.³⁶ Dies betrifft nicht nur Blogs und Websites, sondern insbesondere Webmail-Dienste, die oft als Webspeicherort missbraucht werden, indem wichtige Dateien als E-Mail-Attachment angehängt und an die eigene E-Mail-Adresse versendet werden.³⁷

2.4.3 Datenmigration

Beim Kauf eines neuen Computers stellt sich für den privaten Anwender in der Regel die Frage: Was tun mit den alten Dateien? Hierfür kommen laut Marshall³⁸ drei Lösungsoptionen in Frage:

1. Die Dateien werden auf dem alten Computer belassen und dieser wird aufbewahrt, um zu gegebener Zeit einmal auf die Dateien zurückgreifen zu können;
2. alle Dateien werden aus dem alten Benutzerverzeichnis in das neue Benutzerverzeichnis kopiert;

³⁴ Vgl. Marshall 2008a

³⁵ Vgl. Marshall 2008a

³⁶ Vgl. Williams et al. 2008

³⁷ Vgl. Williams et al. 2008; Marshall 2007, S. 2

³⁸ Vgl. Marshall 2008a

3. die Dateien werden erst bewertet und aussortiert und dann auf den neuen Computer übertragen.

Nach Marshalls Studien wird in der Regel die Kopiermethode ohne Bewertung angewandt, da diese Methode die zeitsparendste ist und keine alte Hardware bereitgehalten werden muss.³⁹ So werden alte Dateien von einem Computer zum nächsten transportiert, ohne Überprüfung, ob die Dateien wirklich zukünftig genutzt werden und ob diese Dateien überhaupt noch mit neuerer Software geöffnet werden können. Auch alte Speichermedien wie Zip-Laufwerke werden aufbewahrt, anstatt die Daten auf neuere Datenträger zu migrieren.⁴⁰ Dies stellt auch Copeland fest:

While participants were mainly motivated to save information for future reference (n=15), when selecting redundant storage, reliability of the storage device was not considered by any of the participants.⁴¹

Zum Problem der als Archivierung missverstandenen Speicherung kommt also noch das Problem der fehlenden Bewertung und Überprüfung von Dateien. Wie weiter oben mit Beagrie zitiert, ist die Verwaltung und Bewertung von digitalen Objekten aufgrund der Masse und der scheinbar günstigeren Retrieval-Methode Suche für die meisten Nutzer zu aufwändig, um sie tatsächlich stetig durchzuführen. Noch ein weiterer Aspekt spielt hierbei eine Rolle: Archivierung ist nicht der Hauptzweck des Handelns, sondern ein Nebeneffekt.⁴² Die Datei wird nur abgelegt, weil sie noch für einen bestimmten Zweck benötigt wird, z. B. Bilddateien, die von einer Kamera auf dem Computer zwischengespeichert werden, um sie von dort für die Präsentation im Internet hochzuladen. Eine weitere Bearbeitung der Datei oder die Anreicherung mit Metadaten findet dann eventuell nur im Internetspeicherort statt, eine Bearbeitung und Anreicherung der lokalen Kopie der Datei wäre erhöhter Aufwand ohne Erfüllung eines dringenden Zwecks. Eine private Archivierungsstrategie, die z. B. besagt, dass sämtliche Bilddateien immer an einem bestimmten Ort gespeichert, eindeutig benannt und mit Metadaten angereichert werden, existiert in der Regel nicht.

Formatmigrationen bei privaten Anwendern sind in Marshalls Studien überhaupt nicht dokumentiert. Marshall sieht die Gründe hierfür im fehlenden Bewusstsein der Nutzer für Formate. Nutzer speichern Dateien in proprietären, durch die jeweilige Software vorgegebenen Formaten, da sie es nicht besser wissen. Solange die entsprechende Software zur Verfügung steht, stellt dies auch kein Problem dar, erst wenn die benötigte Software nicht mehr verfügbar ist und die Dateien nicht mehr geöffnet werden können, wird das Problem bewusst.⁴³ Auch Copeland entdeckt dieses Verhalten bei der von ihr untersuchten Personengruppe:

Participant knowledge of preservation practices related to storage media and file format selection were limited. All 26 participants allowed digital devices and/or Microsoft Office software to determine which formats they used rather than personal preference or knowledge of which formats were best for long-term access, most of the time.⁴⁴

³⁹ Vgl. Marshall 2008a

⁴⁰ Vgl. Marshall et al. 2007, S. 4

⁴¹ Copeland 2011, S. 1295

⁴² Vgl. Marshall 2008b

⁴³ Vgl. Marshall et al. 2007, S. 4f.

⁴⁴ Copeland 2011, S. 1295

Aber noch eine andere Annahme steckt eventuell hinter dieser fehlenden Beschäftigung mit der Obsoleszenz von Formaten bzw. der für die Interpretation benötigten Software, nämlich der Glaube daran, dass ein heute weit verbreitetes Format auch in mehreren Jahren noch gebräuchlich sein wird bzw. die Technik in der Lage sein wird, Lösungen zu schaffen, um weiterhin Zugriff auf Dateien in diesem obsoleten Format zu haben.⁴⁵ Dabei sollten die Nutzer, die bereits mit den ersten Versionen von Microsoft Word Dokumente erstellt haben, bereits die Erfahrung gemacht haben, dass diese Dateien, obwohl es sich um ein weit verbreitetes Format gehandelt hat, mit heutigen Versionen von Word gar nicht oder nur mit Informationsverlust geöffnet werden können.

2.4.4 Bewertung von digitalen Objekten

Ein weiterer untersuchter Aspekt des Archivierungsverhaltens privater Anwender ist der Wert, der einer Datei durch den Nutzer zugesprochen wird.

Hat sich der Nutzer entschieden, eine Datei zu speichern, ist dadurch noch nicht der Wert der Datei für den Nutzer festgelegt. In der klassischen Archivwissenschaft wird der Wert eines Objekts durch Begutachtung festgestellt, wobei laut Cushing dieser Wert in der Regel auch individuell durch den jeweiligen begutachtenden Archivar oder Kurator festgelegt wird und die Bewertung durch verschiedene Archive zu verschiedenen Einschätzungen führen kann, hier also auch keine automatisierte Bewertung stattfindet.⁴⁶ Cathy Marshall behauptet, dass der Wert einer Datei für den Nutzer nur selten explizit benannt werden kann; vielmehr wird er durch die folgenden Indikatoren angezeigt:

- Nutzungshäufigkeit
- Replikationshäufigkeit
- Zusätzliche Speicherung als E-Mail-Attachement
- Vorgenommene Dateibenennung
- Teilen mit anderen⁴⁷

Dies impliziert zum Beispiel, dass gerade die Archivierung im klassischen Sinne, also ohne Ad-hoc-Zugriff für den Nutzer, ein Indikator dafür wäre, dass die Datei als wenig wertvoll betrachtet wird.

Williams et al. bestätigen in ihrer Studie die von Marshall genannten Indikatoren. Der generelle Wert einer selbsterstellten Datei in der Wahrnehmung des Nutzers ist laut ihrer Untersuchung durch den Aufwand bestimmt, der für ihre Erstellung betrieben wurde.⁴⁸ Marshall nennt dies den „emotional impact“⁴⁹. Auch Copeland stellt diesen „emotional impact“ fest, sieht ihn aber eher in der Interaktion mit anderen begründet:

In the everyday context, personal information is most valued for communication with other people. In the long-term context, personal information is most valued for the connection to emotions associated with friends and family that is motivated by the documentation of life events.⁵⁰

⁴⁵ Vgl. Marshall 2008a

⁴⁶ Vgl. Cushing 2010, S. 309

⁴⁷ Vgl. Marshall 2008a

⁴⁸ Vgl. Williams et al. 2008

⁴⁹ Marshall 2008b

⁵⁰ Copeland 2011, S. 1296

Würden die oben genannten Indikatoren stetig zur Bewertung angewendet, könnte man sie tatsächlich als Kriterien festschreiben. In der Regel erfolgt die Anwendung aber nur selten; stattdessen werden im Rahmen von „Löschfeldzügen“ relativ wahllos Dateien gelöscht werden, z. B. weil tatsächlich Speicherplatz freigegeben werden soll oder weil der gefrustete Anwender dem digitalen Heuhaufen Herr werden möchte.⁵¹ Eine Strategie ist hierbei nicht vorhanden.

Generell scheint das Löschen von einmal gespeicherten Dateien die für den Nutzer am schwierigsten zu bewältigende Aufgabe zu sein, egal ob es sich dabei um selbsterstellte Dateien handelt oder nicht. Hardwarecrashes werden in diesem Zusammenhang sogar positiv gesehen, da man auf diese Weise ohne eigenes Zutun vom digitalen Ballast befreit wurde:

Periodic loss is regarded as inevitable and in principle as a way of controlling clutter, much like forest fires may be used to clear brush to support the health of the overall ecosystem.⁵²

Das Nutzerverhalten hinsichtlich der Bewertung ist also durch relatives Schwarz-Weiß-Denken geprägt: Auf der einen Seite wird alles gespeichert und damit alles als gleichwertig angesehen, auf der anderen Seite wird der Verlust aller Dateien als Katharsis betrachtet.

Williams et al. führen aber noch eine alternative Interpretation der Bewertung von digitalen Objekten an:

Indeed, some respondents look at their archive as a reflection of their life's work, and keep items of no further practical value.⁵³

In diesem Fall werden Dateien nicht einzeln bewertet, sondern sie werden in ihrer Gesamtheit als wertvoll betrachtet, da sie das Leben des jeweiligen Menschen dokumentieren. Amber Cushing leitet hieraus die folgende Archivierungsstrategie ab:

[...] if individuals were to privilege special life events over others, digital belongings related to those specific life events could be kept and digital belongings related to not so important events could be discarded.⁵⁴

Marshall hingegen analysiert vier verschiedene Wertigkeiten, die auf die angesammelte Datenmenge angewendet werden können:⁵⁵

1. Bekannte Daten von hohem Wert: Hierbei handelt es sich in der Regel um einige wenige Dateien, die problemlos identifiziert werden können.
2. Teilsammlungen von mittlerem Wert: Diese sind in der Regel umfangreich und haben unterschiedlichen Wert, der sich mit der Zeit verändern kann.
3. Teilsammlungen mit niedrigem Wert: Diese Sammlungen enthalten mehrheitlich digitalen Ballast, also Daten, die nicht im direkten Gebrauch sind und deren Bewertung daran scheitert, dass diese Aufgabe als zu aufwändig betrachtet wird. Darunter können sich daher aber auch einige wenige Dateien von Wert befinden, die aber erst identifiziert werden müssen.

⁵¹ Vgl. Marshall 2008a

⁵² Marshall 2008a

⁵³ Williams et al. 2008

⁵⁴ Cushing 2010, S. 308

⁵⁵ Vgl. Marshall 2011, S. 103f.

4. Belastende Dateien: hierbei handelt es sich um Daten, die man tatsächlich loswerden möchte, in der Regel aus emotionalen oder rechtlichen Beweggründen. Das Problem bei diesen Daten ist also nicht die Erhaltung, sondern die gründliche und unwiederbringliche Löschung.

Diese Werte zu bestimmen stellt dabei die größte Herausforderung und den größten Aufwand für den Anwender dar, weshalb diese Dateien in der Regel so belassen werden wie sie sind.

2.4.5 Zusammenfassung

Das in den Studien von Marshall, Williams et al. und Copeland beobachtete Archivierungsverhalten zeigt eine Fülle von Verhaltensweisen, die aber individuell unterschiedlich ausfallen können. Williams et al. sehen

[...] significant differences in

- Methods and places of storage;
- Familiarity and expertise with hardware and software;
- Understanding of the meaning of a 'personal digital collection' (respondents' own views of this concept formed part of the project, and so it was not incumbent on the researchers to provide more than a general explanation);
- Individual perceptions of what and especially how much is worth keeping (as is the case with conventional archives too);
- Relative values attached to digital and non-digital items.⁵⁶

Demnach sind die ermittelten Probleme bei der Archivierung sowohl technischer Art (Speichermethoden und -orte, Erfahrung mit Hardware und Software) als auch konzeptueller Art (Was sind persönliche digitale Kollektionen, was und wieviel sollte gespeichert werden, wie kann der Wert bestimmt werden?).

Die Dokumentation der zitierten Studien ist leider sehr uneinheitlich. Marshall und Williams et al. bieten nur Zusammenfassungen der durchgeführten Untersuchungen; lediglich Copeland präsentiert eine wissenschaftliche Dokumentation ihrer durchgeführten Untersuchung, komplett mit Vorstellung der Methode, der Teilnehmer und der statistischen Auswertung.

Alle Studien untersuchen jeweils das Archivierungsverhalten zu einem bestimmten Zeitpunkt. Langzeitstudien, die auch Änderungen in den Verhaltensweisen dokumentieren könnten, fehlen bisher. Dies bemerkt auch Copeland:

A long-term study of PIM is needed to understand how the access issues affect an individual's ability to access his or her own content. These issues would include media durability, system and software changes, and organizational structures that provide access to the content of personal files.⁵⁷

2.5 Konzepte und Modelle

Die weiter oben skizzierte wissenschaftliche Diskussion zum Thema Personal Digital Archiving hat nur wenige Konzepte, Modelle und Handlungsempfehlungen hervorge-

⁵⁶ Williams et al. 2008

⁵⁷ Copeland 2011, S. 1298

bracht; die jeweiligen Ansätze und die Gründe für den Mangel an ebensolchen werden im Folgenden vorgestellt.

Wie im vorherigen Abschnitt erläutert wurde, ist die Verwaltung und Archivierung persönlicher digitaler Daten geprägt durch wohlwollende Vernachlässigung, Ansammlung großer Datenmengen an verschiedenen, teilweise cloudbasierten und passwortgeschützten Speicherorten und der stoischen Wahrnehmung von Datenverlust als reinigendes Ereignis. Größtes Hindernis der Verwaltung und Archivierung ist der zeitliche Aufwand für die Pflege der Daten, der in der Wahrnehmung der Nutzer in keiner Relation zum Ergebnis steht. Stattdessen vertrauen die Anwender darauf, dass die IT zu gegebener Zeit entsprechende Lösungen bereitstellen wird, um digitale Daten vor dem Verlust zu retten.

Die Digital-Lives-Studie kann unter den befragten Nutzern keine eindeutige Archivierungsstrategie ausmachen, sondern findet in den Interviews mit den privaten Anwendern „a whole range of ‚digital‘ behaviours“⁵⁸ vor. Bei den Interviewten handelt es sich um eine heterogene Gruppe, die sowohl in unterschiedlichem Maße mit dem Computer arbeitet, die technischen Zusammenhänge versteht (z. B. welche Daten lokal gespeichert werden und welche Daten online gespeichert sind) als auch unterschiedliche Auffassungen vertritt, welche und wie viele Daten gespeichert werden und welchen Wert diese digitalen Daten für sie haben.⁵⁹

Diese Heterogenität macht es schwer, auf Basis dieser Erkenntnisse eine eindeutige Handlungsstrategie zu formulieren. Entweder muss ein kleinster gemeinsamer Nenner gefunden werden, auf dem eine Best-Practice-Methode aufbauen kann oder es muss eine Generalisierung der vorgefundenen Handlungsweisen vorgenommen werden und darauf aufgebaut werden. Eine effektive Archivierungsstrategie sollte auf dem derzeit praktizierten Verhalten der Anwender aufbauen. Dieses ist nicht mit den Best-Practice-Empfehlungen kompatibel, wie sie z. B. in der professionellen Langzeitarchivierung erarbeitet werden. Hochkomplexe theoretische Modelle, wie sie in der institutionellen Langzeitarchivierungsforschung beschrieben werden, sind für den individuellen Endnutzer vermutlich nicht anwendbar.

Ein zentraler Speicherort für alle persönlichen Daten wäre zwar die praktikabelste und übersichtlichste Lösung, entspricht aber nicht dem Nutzerverhalten.⁶⁰ Außerdem hat die verteilte Speicherung gegenüber der zentralen Speicherung natürlich auch den Vorteil, dass niemals alle Daten gleichzeitig verloren gehen. Marshall plädiert daher für eine Art zentralen Katalog, der auf Metadatenebene sämtliche Dateien mit ihren Speicherorten und Eigenschaften registriert. Alerting-Funktionen weisen auf auslaufende Services oder bedrohte Formate hin. Dies entspricht ihrer Ansicht nach am ehesten dem analogen Archivierungsverhalten „benign neglect“: Daten werden ohne ständige Verwaltung oder Kontrolle gespeichert, nur wenn es wirklich nötig ist, werden entsprechende Handlungen vorgenommen.⁶¹

Da die bereits erwähnte Digital-Lives-Studie vor allem darauf ausgerichtet war, persönliche digitale Sammlungen in institutionelle Repositorien aufnehmen zu können, werden modellartig Module und Aktivitäten definiert, die für die erleichterte Übernahme der Kollektionen erforderlich sind.⁶² Für Personal Digital Archiving an sich werden keine Modelle entworfen.

⁵⁸ Williams et al. 2008

⁵⁹ Vgl. Williams et al. 2008

⁶⁰ Vgl. Marshall 2011, S. 104ff.

⁶¹ Vgl. Marshall 2011, S.106

⁶² Vgl. John et al. 2010, S. 184

2.6 Technische Unterstützung

Im vorigen Abschnitt wurde bereits eine Theorie vorgestellt, die eine technische Lösung für die Langzeitarchivierung persönlicher digitaler Daten skizziert. Derzeit sind bereits Lösungen verfügbar, die auf unterschiedlichem Niveau hinsichtlich Archivierungsqualität und Nutzerfreundlichkeit angesiedelt sind. Diese technischen Lösungen sollen im Folgenden kurz vorgestellt werden.

2.6.1 Backup-Lösungen

Neben der Möglichkeit des manuellen Backups von Dateien verfügen viele Betriebssysteme derzeit bereits über Backuproutinen. Durch die Erstellung von Systemabbildern bzw. Images ist es z. B. unter Microsoft Windows 7 möglich, den Computer so wiederherzustellen, wie er zu einem bestimmten Zeitpunkt installiert und konfiguriert war, inklusive der Dateien auf der Festplatte. Apple bietet mit Time Machine unter Mac OS X eine automatische Lösung für die Sicherung sämtlicher Konfigurationen und Daten, in der auch im laufenden Betrieb gelöschte Dateien bzw. Dateiversionen gesucht und wiederhergestellt werden können.

Da es sich bei beiden Lösungen um reine Backupfunktionen handelt, ermöglichen sie nur Bitstream Preservation, also die Bewahrung des ursprünglichen Datenstroms. Für den privaten Nutzer bringen diese Funktionen aber den Vorteil mit sich, das ihm überhaupt vollumfängliche Sicherungen seiner Daten möglich sind, die er dann auf externe Speichermedien kopieren kann. Wie bereits weiter oben erwähnt, ist das Erstellen eines Backups die einfacher zu bewältigende Seite des Prozesses; ob die Erstellung des Backups tatsächlich funktioniert hat, zeigt sich erst, wenn das Backup wiedereingespielt werden muss, ein Prozess mit dem die wenigstens Nutzer Erfahrung haben und der den tatsächlichen Knackpunkt beim Backup darstellt. Denn was nützt das Backup, wenn es nicht eingespielt werden kann, sei es bedingt durch Inkompatibilität oder Inkompetenz.

2.6.2 MyLifeBits

MyLifeBits⁶³ ist ein Forschungsprojekt von Microsoft Research, in dem der Wissenschaftler Gordon Bell im Selbstversuch ein digitales Gedächtnis seines Lebens aufbaut. Hierzu wurden in einer ersten Phase sämtliche Arbeitspapiere und private Dokumente digitalisiert, einschließlich Nachrichten auf dem Anrufbeantworter.⁶⁴ Sämtliche neu hinzukommenden Dokumente und Sprachdaten werden automatisch der MyLifeBits-Datenbank hinzugefügt und für die Suche nach Volltexten und Metadaten aufbereitet. Gordon Bell trägt ständig eine sogenannte SenseCam, die je nach verändertem Lichteinfall und dem damit zu erwartenden veränderten Ambiente Fotos aufnimmt, um diese Veränderung zu dokumentieren.⁶⁵

Wie aus dieser kurzen Beschreibung bereits zu entnehmen ist, geht das Ziel des MyLifeBits-Projekts weit über das hinaus, was in dieser Arbeit als Ziel von Personal Digital Archiving definiert wurde, nämlich die (Langzeit-)Archivierung persönlicher digitaler Dateien. Vielmehr geht es in diesem Projekt tatsächlich um die Erweiterung des eigenen Gedächtnisses um das digitale Gedächtnis, das so schon in den 40er Jahren des letzten Jahrhunderts von Vannevar Bush angedacht wurde.⁶⁶ In Zeiten von NSA-

⁶³ Siehe Microsoft Research 2014

⁶⁴ Vgl. Bell und Gemmel 2007, S. 88

⁶⁵ Vgl. Bell und Gemmel 2007, S. 90

⁶⁶ Vgl. Vannevar Bush 1945

Spionage und Google-Datensammlung wirkt es jedoch ein wenig befremdlich, dass man selber jede Aussage und jeden Eindruck innerhalb seines Lebens speichert und diese Dinge natürlich nur zu guten Zwecken wie z. B. der Überwachung von Gesundheitsdaten⁶⁷ genutzt werden sollen.

Auch sollte man beachten, dass der zitierte Artikel aus dem Jahr 2007 stammt, also aus einer Zeit vor dem Boom von Smartphones, Twitter, Facebook und Selfies. Die Dokumentation jedes sich bietenden Eindrucks haben die Endanwender mithilfe dieser Tools aktiv selber übernommen – die Archivierung und Verknüpfung dieser Daten betreiben Dritte unter dem Stichwort Big Data. Als Softwarelösung für Personal Digital Archiving scheidet dieser Ansatz jedenfalls aus.

2.6.3 HOPPLA

HOPPLA⁶⁸ ist eine von Forschern der TU Wien konzipierte und teilweise auch umgesetzte Lösung für die private Langzeitarchivierung. Ziel der Entwicklung war es, die Vorgaben der institutionellen Langzeitarchivierung wie z. B. das OAIS-Modell in das Softwaredesign einzubeziehen, die Software aber für den unerfahrenen Anwender einfach zu halten.⁶⁹ Dies sollte dadurch gelöst werden, dass ähnlich wie bei Antivirusprogrammen mittels einer Client-Server-Architektur die erforderlichen Schritte und Aufgaben für die Langzeitarchivierung nicht durch den lokal installierten Client entschieden und durchgeführt werden, sondern basierend auf den Metadaten der Sammlung werden die Dateien vom Server per Update-Service regelmäßig überprüft und bei Bedarf entsprechende Migrationsaktionen vorgeschlagen bzw. die jeweilige Strategie angepasst.⁷⁰ Bei der tatsächlichen Durchführung wird dann wieder auf lokal installierte Hilfsprogramme wie z. B. ImageMagick für die Formatkonvertierung zurückgegriffen.⁷¹ Durch den Abruf externer Expertise auf Grundlage der ermittelten Metadaten per Update Service soll der Anwender relativ wenig mit aktiven Entscheidungen hinsichtlich der Langzeitarchivierung belastet werden.⁷²

Auffällig sind die Ähnlichkeiten zu dem im vorigen Kapitel skizzierten Lösungsvorschlag von Cathy Marshall, der ein zentrales Repository für Metadaten der zu archivierenden Dateien vorsieht, über das die jeweiligen Aktionen gesteuert werden. Der Update-Service-Ansatz sorgt dafür, dass immer auf die aktuellsten Erkenntnisse zurückgegriffen werden kann.

Allerdings konnten bislang nur bestimmte Teile des Lösungskonzepts tatsächlich umgesetzt werden. Auch scheint sich die Zielgruppe verschoben zu haben; in neueren Aufsätzen wird der private Anwender weitgehend ausgeklammert, angesprochen werden nur noch die kleineren Unternehmen. Eventuell liegt hier also auch eine Änderung des Geschäftsmodells vor – wenn man davon bei solchen Forschungsprojekten überhaupt ausgehen kann: Um einen Webservice wie den skizzierten in qualitativ hochwertiger Form anbieten zu können, werden wahrscheinlich entsprechende zahlende Kunden benötigt. Allerdings ist die letzte Version 2.1 der HOPPLA-Software aus dem Jahr 2010; seither sind auch keine weiteren Publikationen recherchierbar. Daher scheint diese Lösung mit der Einstellung des EU-Projekts PLANETS nicht weiter entwickelt wor-

⁶⁷ Vgl. Bell und Gemmel 2007, S. 90

⁶⁸ Siehe TU Wien

⁶⁹ Vgl. Strodl et al. 2008, S. 1ff.

⁷⁰ Vgl. Strodl et al. 2008, S. 4ff.

⁷¹ Vgl. Strodl et al. 2010, S. 128

⁷² Vgl. Strodl et al. 2010, S. 129

den zu sein. Somit steht auch mit HOPPLA keine Lösung für den privaten Endanwender bereit.

2.6.4 Zusammenfassung

Aufgrund der beschriebenen Probleme bei der Anwendung der vorhandenen Softwarelösungen durch Privatanwender scheint es derzeit keine bessere Alternative zu reinen Backup-Lösungen zu geben. Um aber Backups effizienter zu gestalten, bedarf es der Vermittlung und Schulung von Best-Practice-Methoden, die dem Nutzer die Archivierung erleichtern und positive Effekte auf die spätere Wiederauffindbarkeit der Daten haben. Entsprechende Ansätze werden im folgenden Kapitel vorgestellt.

2.7 Informations- und Schulungsangebote

Die Vermittlung und Schulung von Techniken zur privaten Archivierung digitaler Objekte findet auf Grund des fehlenden Problembewusstseins so gut wie gar nicht statt. Überwiegend sind Anleitungen zur Archivierung digitaler Fotos in den Ratgeberteilern von Magazinen und Zeitschriften zu finden, die sich z. B. im Rahmen eines Dossiers einmalig mit dem Thema beschäftigen. Dabei hängt die Qualität der Archivierungstipps natürlich sehr von der Recherchetiefe und dem Fachwissen des Autors ab. Auch das bereits vorgestellte Portal *Digit* hält einen Link zu einem entsprechenden Beitrag des WDR-Verbrauchermagazins *markt* bereit, der aber nicht mehr online ist.

Professionelle Materialien zur privaten Archivierung digitaler Objekte in deutscher Sprache konnten nicht ermittelt werden. Anbieter umfangreicher und auf verschiedene Objekttypen zugeschnittener Informationsangebote im englischsprachigen Raum ist die Library of Congress. Im Rahmen des National Digital Information Infrastructure and Preservation Program (NDIIPP) bietet sie auf ihrer Website zu Personal Digital Archiving⁷³ einfache Handreichungen an, die für jeden Objekttyp auch durch den Laien durchzuführende Langzeitarchivierungsmaßnahmen erläutern. Videos, ein Quiz und das Blog *The Signal* liefern zusätzliche Informationen zu Personal Digital Archiving. Für interessierte Bibliotheken steht ein Personal Archiving Day Kit⁷⁴ bereit, mit dem Bibliotheken bei der Organisation eigener Veranstaltungen zum Thema unterstützt werden sollen.

Bei der Vermittlung müssen daher auch zwei Zielgruppen unterschieden werden: Zum einen der private Nutzer, der Anleitung für die Archivierung seiner digitalen Daten erhalten möchte, zum anderen das fachlich geeignete Personal, das dem Nutzer die entsprechende Anleitung geben kann. Das InterPARES-Projekt⁷⁵ hat daher in seiner zweiten Phase Guidelines sowohl für die Produzenten von digitalen Objekten, als auch für das archivierende Personal erstellt. Diese Anleitungen wurden in Englisch, Französisch, Spanisch, Portugiesisch und Katalanisch erstellt und sind auf der Website als PDF abrufbar.⁷⁶

Als Ergebnis des PARADIGM-Projekts wurde nur für die Erzeuger digitaler Objekte eine Anleitung erstellt. Diese ist Teil des PARADIGM-Workbooks und kann online abgerufen werden.⁷⁷

⁷³ Library of Congress o.J.a

⁷⁴ Library of Congress o.J.b

⁷⁵ Siehe InterPARES Project o.J.

⁷⁶ Siehe InterPARES 2 Project o.J.

⁷⁷ Siehe PARADIGM 2008a

Informationsangebote zu Personal Digital Archiving sind also durchaus vorhanden, wobei die Seiten der Library of Congress wahrscheinlich besser recherchierbar sind als die Ergebnisse der beiden Projekte; Qualität, Aufbereitung und Zielgruppen werden in Kapitel 5.1 näher untersucht.

2.8 Zusammenfassung

Die bisherigen Untersuchungen zum Archivierungsverhalten von Einzelpersonen zeigen uneinheitliche Vorgehensweisen, Schwierigkeiten bei der Bewertung der gespeicherten Dateien und Defizite hinsichtlich der Formate und Speichermethoden. Verteilte Speicherung und große Datenmengen tragen zur Unauffindbarkeit von Dateien bei. Softwareprodukte lösen die beobachteten Probleme nur teilweise. Informations- und Schulungsangebote existieren nur in englischer Sprache und sind teilweise schwer recherchierbar. Um einen Vergleich dieser Ergebnisse mit den Konzepten und Methoden der institutionellen Langzeitarchivierung vornehmen zu können, werden diese im folgenden Kapitel vorgestellt.

3 INSTITUTIONELLE LANGZEICTARCHIVIERUNG

3.1 Begriffsdefinition

Der Begriff Langzeitarchivierung scheint auf den ersten Blick eine Tautologie zu sein, da Archivierung ja in der Regel impliziert, dass die archivierten Objekte für einen längeren Zeitraum vorgehalten werden. Langzeitarchivierung digitaler Objekte unterscheidet sich aber vom normalen Archivierungsbegriff dadurch, dass es keinen bestimmten Zeithorizont gibt, bis zu dem das Archivgut verfügbar sein muss. Langzeitarchivierung bedeutet, dass die archivierten Objekten so lange wie möglich erhalten werden, unabhängig von äußeren Einflüssen:

„Langzeit“ bedeutet für die Bestandserhaltung digitaler Ressourcen nicht die Abgabe einer Garantieerklärung über fünf oder fünfzig Jahre, sondern die verantwortliche Entwicklung von Strategien, die den beständigen, vom Informationsmarkt verursachten Wandel bewältigen können.⁷⁸

Ebenfalls in Gebrauch ist der Begriff der Langzeitverfügbarkeit. Im Gegensatz zur Archivierung, die impliziert, dass Objekte nicht zugänglich sind bzw. nicht für die Nutzung aufbereitet werden, ist bei der Langzeitarchivierung Zugriff und Nutzung eingeplant. Eventuell kann man die beiden Begriffe folgendermaßen in Relation setzen: Langzeitarchivierung ist erforderlich, um die Langzeitverfügbarkeit sicherstellen zu können.

3.2 Problemstellung

Ziel der Langzeitarchivierung digitaler Objekte ist die Langzeitverfügbarkeit dieser Objekte für künftige Nutzergenerationen. Auf Grund der Schnelllebigkeit von Datenformaten und Datenträgern sowie der zum Zugang benötigten Software sind hierzu wiederkehrende Eingriffe wie die erneute Speicherung auf neuen Datenträgern sowie die Umwandlung in neue Formate unter Beibehaltung der das Objekt ausmachenden Eigenschaften nötig, um sicherzustellen, dass folgende Generationen auf diese Daten und Objekte ohne Informationsverlust zugreifen können. Probleme hierbei sind die Menge der zu archivierenden Daten und ihre Reduzierung durch entsprechende Auswahlprozesse, geeignete Datenträger und Formate, Metadaten, Finanzierung und Organisation sowie rechtliche Fragestellungen. Damit kommen bei der digitalen Langzeitarchivierung alle Themengebiete zusammen, die auch beim analogen Bestandsaufbau und der analogen Bestandserhaltung eine Rolle spielen.

3.3 Wissenschaftliche Diskussion

Die wissenschaftliche Diskussion der digitalen Langzeitarchivierung behandelt im Grunde ein Thema, das die Wissenschaftskommunikation selber betrifft. Wissenschaftliche Fachkommunikation, Datenerhebung und Publikation findet mittlerweile fast ausschließlich digital statt. Um diese Erkenntnisse und Ergebnisse für aktuelle und kommende Forschergenerationen zu erhalten, müssen entsprechende Maßnahmen getroffen werden, um die Langzeitverfügbarkeit sicherzustellen. Diese Maßnahmen werden unter

⁷⁸ Schwens und Liegmann 2004

dem Begriff der digitalen Langzeitarchivierung zusammengefasst. Dabei geht es natürlich nicht nur um die Archivierung von Forschungsdaten, die sich mittlerweile zu einem eigenen Forschungsschwerpunkt entwickelt hat, sondern generell um die Archivierung digitaler Daten durch Bibliotheken, Museen und Archive, also die traditionellen Gedächtnisorganisationen.

In vielen Ländern, z. B. in den USA, Australien und Dänemark sind die Nationalbibliotheken mit der digitalen Langzeitarchivierung des nationalen Kulturguts betraut, seien es elektronische Dokumente oder Webseiten. In Deutschland ist die Deutsche Nationalbibliothek per Gesetz zur Archivierung auch elektronisch abgelieferter Pflichtexemplare und Webpublikationen verpflichtet.⁷⁹

Auf Grund dieser Verpflichtung sind auf internationaler und nationaler Ebene vielfältige Initiativen und Netzwerke entstanden, die Fragen der institutionellen digitalen Langzeitarchivierung diskutieren. In Deutschland ist dies das Kompetenznetzwerk nestor⁸⁰, dessen Geschäftsstelle an der Deutschen Nationalbibliothek angesiedelt ist. In verschiedenen Arbeitsgruppen werden hier Themen wie digitale Bestandserhaltung, Kosten der digitalen Langzeitarchivierung und rechtliche Fragen behandelt. Besetzt sind die Arbeitsgruppen mit Experten unterschiedlicher deutscher Institutionen, die sich um eine gemeinsame Klärung der jeweiligen thematischen Fragen und um gemeinsame Standards und Vorgehensweisen bemühen. Hieraus sind bereits einige Publikationen zum Thema digitale Langzeitarchivierung entstanden; die bekannteste und zentrale Publikation ist das nestor-Handbuch, das aktuell in der Version 2.3 aus dem Jahr 2010 vorliegt und über die nestor-Website abrufbar ist.⁸¹

International haben sich ähnliche Initiativen gebildet, die in der Regel ebenfalls bei den jeweiligen Nationalbibliotheken angesiedelt sind. In den USA ist dies das bereits erwähnte NDIIPP, aus dem weitere Initiativen wie die National Digital Stewardship Alliance hervorgegangen sind, die sich ähnlich wie nestor um Austausch und gemeinsame Forschung und Standards bemüht, sowie die Initiative Digital Preservation Outreach and Education, die sich um Schulung und Qualifizierung kümmert. Die British Library fördert die Digital Preservation Coalition, die ebenfalls als Verbund von Experten verschiedener UK-Bibliotheken Probleme der digitalen Langzeitarchivierung behandelt.

Neben diesen jeweils nationalen Initiativen wurden insbesondere auf europäischer Ebene verschiedene länderübergreifende Projekte initiiert. Bereits kurz erwähnt wurde das PLANETS-Projekt, das sich vorwiegend mit der Erstellung von Software für bestimmte Prozesse der digitalen Langzeitarchivierung auseinandersetzte. Dieses ist mittlerweile in die Open Planets Foundation übergegangen, setzt sich aber weiterhin mit dem Thema Softwareentwicklung für zentrale Anforderungen der digitalen Langzeitarchivierung auseinander, z. B. der Entwicklung von Tools für die Formaterkennung. Nachfolger des PLANETS-Projekts ist SCAPE, das sich mit der Entwicklung skalierbarer Open-Source-Dienste für die digitale Langzeitarchivierung beschäftigt. Eine Übersicht weiterer Projekte und Initiativen bietet das nestor-Wiki.⁸²

Die vorgestellten Projekte zeichnen sich in erster Linie durch technische Fragestellungen aus. Dies ist dadurch bedingt, dass aufgrund der enormen Datenmenge Tools zur Automatisierung bestimmter Vorgänge dringend benötigt werden. Aber auch andere

⁷⁹ Siehe Gesetz über die Deutsche Nationalbibliothek 2006

⁸⁰ Network of Expertise in long-term **S**torage and availability of digital **R**esources in Germany; siehe nestor o.J.b

⁸¹ Siehe Neuroth, Heike et al. 2010

⁸² Siehe nestor o.J.a

Aspekte, z. B. die Kosten der Langzeitarchivierung werden in entsprechenden Projekten erforscht, wie z. B. im europäischen 4C-Projekt.

Zusammengefasst beschäftigt sich die wissenschaftliche Diskussion vorwiegend mit technischen Aspekten der digitalen Langzeitarchivierung, wobei Kooperationen sowohl auf nationaler Ebene wie auch auf europäischer bzw. internationaler Ebene die Norm darstellen. Dadurch sind eine gute Vernetzung der beteiligten Partner und eine gemeinsame Herangehensweise gewährleistet.

3.4 Archivierungsverhalten

Das Archivierungsverhalten auf institutioneller Ebene ist durch verschiedene Vorgaben und Richtlinien bestimmt, um eine nachhaltige und vertrauenswürdige Archivierung zu gewährleisten. Diese Richtlinien definieren das Kollektionsprofil, die Bestimmung des Werts eines elektronischen Objekts für zukünftige Nutzergenerationen sowie die durchzuführenden Handlungen im Rahmen der Archivierungsstrategie. Grundlegend hierfür ist die Preservation Policy.

3.4.1 Preservation Policy

Bei einer Preservation Policy handelt es sich um eine Richtlinie, die den äußeren verbindlichen Rahmen sämtlicher Langzeitarchivierungshandlungen bildet. Sie geht dabei nicht so weit ins Detail, dass sie spezielle Arbeitsanweisungen enthält, definiert aber Aufgaben, Tätigkeiten und Pflichten in der Handhabung der zu archivierenden Objekte. Es wird zwischen einer nationalen Policy⁸³, die sich auf den nationalen Archivierungsrahmen bezieht und institutionellen Policies⁸⁴ unterschieden, die jede Institution, die Langzeitarchivierung betreibt, für sich selber erstellen muss. Im Folgenden soll insbesondere auf Sinn und Zweck institutioneller Policies eingegangen werden.

Eine treffende Definition des Zwecks von Preservation Policies gibt die nestor-Arbeitsgruppe Policy:

Policy-Dokumente dienen der Selbstvergewisserung und der Selbstverpflichtung. Sie legen langfristig wirksame strategische und organisatorische Grundelemente eines digitalen Langzeitarchivs offen und tragen zum allgemeinen Vertrauensgewinn bei. In diesem Sinne helfen Policies, digitale Informationen von gestern und heute verlässlich zu erhalten und für Nutzer von morgen zu bewahren.⁸⁵

Institutionelle Preservation Policies haben also in der Regel zwei verschiedene Zielgruppen: Intern bilden sie die Grundlage sämtlicher Tätigkeiten für die mit der Langzeitarchivierung betrauten Mitarbeiter der Institution sowie für das Management. Die Mitarbeiter der Institution verfügen durch die Policy über eine verlässliche interne Richtlinie, an der sie ihre tägliche Arbeit ausrichten können, das Management muss festgeschriebene organisatorische Ressourcen bereithalten sowie sich zu Betrieb, Finanzierung und Weiterentwicklung bekennen. Extern klärt die Preservation Policy Nutzer und Produzenten digitaler Objekte darüber auf, dass ihre Daten nach einer bestimmten Vorgehensweise archiviert werden und sorgt dadurch für Transparenz.⁸⁶ Daher ist es auch wichtig, dass die aktuellen und zukünftigen Zielgruppen der Archivierungsbemü-

⁸³ Vgl. Strathmann 2010, Kap. 3:3

⁸⁴ Vgl. Strathmann 2010, Kap. 3:6

⁸⁵ nestor-Arbeitsgruppe Policy 2014, S. 2

⁸⁶ Vgl. nestor-Arbeitsgruppe Policy 2014, S. 4

hungen benannt werden (siehe hierzu Kapitel 3.5.2) und darauf eingegangen wird, wie die Nutzer Zugang zu den archivierten Objekten erhalten.⁸⁷

Im Leitfaden der nestor-Arbeitsgruppe Policy wird noch einmal explizit betont, dass organisatorisch im Voraus geklärt sein muss, wie das Archiv dauerhaft finanziert wird und dies auch in der Policy dokumentiert wird.⁸⁸ Besteht keine dauerhafte Finanzierung, ist das Archiv auch nicht unbedingt als vertrauenswürdig einzustufen, da nicht garantiert ist, dass bei Wegfall der Finanzierung der Zugriff der Nutzer auf die archivierten Daten weiter gegeben ist. Für die interne Organisation ist es wichtig, dass Mitarbeiterstellen mit entsprechenden Aufgaben, fachliche und technische Zuständigkeiten sowie die Hardwareausstattung festgeschrieben werden.⁸⁹

Eine Policy wird in der Regel auch sehr genau darauf eingehen, dass die Maßnahmen und Abläufe bei der Handhabung der Daten sowie Änderungen an den Daten entsprechend dokumentiert werden und z. B. mit Hilfe von Prüfsummen die Datenintegrität gewährleistet wird.⁹⁰ Auch dies dient der Vertrauenswürdigkeit des Archivs und macht alle Vorgänge nachvollziehbar und transparent. Sammelrichtlinien werden in der Regel nicht in der Preservation Policy festgelegt; diese ergeben sich entweder aus dem generellen Sammlungsprofil der Einrichtung oder werden gesondert in einer Collection Policy definiert.⁹¹

3.4.2 Archivierungsstrategien

Um digitale Objekte dauerhaft zugänglich zu machen, muss sichergestellt werden, dass potenzielle Nutzer weiterhin mittels Softwarelösungen wie z. B. einem Viewer für Bilddateien oder einem Editor für Textdateien auf die Objekte zugreifen können. Hierzu müssen erst einmal die Originaldaten vor dem Verlust durch Hardwareschäden gesichert werden. Dies erfolgt in der Regel mittels verteilter Speicherung bzw. durch die Einrichtung entsprechender Backup-Systeme. Änderungen an den archivierten Dateien werden dabei nicht vorgenommen. Diese Grundvoraussetzung nennt man Bitstream Preservation; dabei wird nur der reine Datenstrom bewahrt und keine Manipulation der gespeicherten Dateien vorgenommen.

Wie bereits erwähnt, veralten Formate und die sie interpretierenden Softwarelösungen immer schneller. Um heute erstellte Dateien in mehr als 15 Jahren noch nutzen zu können, muss entweder im Rahmen einer Migration die Datei in ein neues Format umgewandelt werden oder durch Emulation die notwendige Software- und Systemumgebung nachgebildet werden. Neben den zu erwartenden Kosten – bei der Migration müssen fortlaufende Kosten eingerechnet werden, bei der Emulation punktuelle Kosten für die Programmierung der jeweiligen Umgebung – unterscheiden sich die beiden Strategien noch in weiteren Gesichtspunkten, die im Folgenden erläutert werden.

Bei der Migration von Daten zum Zwecke der digitalen Langzeitarchivierung werden digitale Objekte in neue Formate übertragen, wenn sich abzeichnet, dass das Format, in dem die Daten originär bzw. aktuell vorliegen, in naher Zukunft nicht mehr durch entsprechende Softwarelösungen unterstützt wird. Hierfür ist es von Vorteil, wenn sowohl das Ausgangs- als auch das Zielformat in Form eines offenen Standards entsprechend

⁸⁷ Vgl. nestor-Arbeitsgruppe Policy 2014, S. 7

⁸⁸ Vgl. nestor-Arbeitsgruppe Policy 2014, S. 8

⁸⁹ Vgl. nestor-Arbeitsgruppe Policy 2014, S. 9

⁹⁰ Vgl. nestor-Arbeitsgruppe Policy 2014, S. 10

⁹¹ Vgl. GESIS 2013

gut dokumentiert sind.⁹² Dadurch können Informationsverluste vermieden bzw. rechtzeitig erkannt werden, wohingegen bei proprietären Formaten ein Verlust von Informationen unbemerkt bleiben kann, da die entsprechende Spezifikation gegebenenfalls nicht vorhanden ist. Technische Objektmetadaten, die mit entsprechenden Formaterkennungstools (siehe Kapitel 3.6.1) vor und nach der Migration abgerufen werden können, geben Aufschluss über die zum Objekt gehörenden Informationen, z. B. bei Bilddateien die Auflösung, Bildmaße, ISO-Belichtung, Scanner- oder Kamerahersteller etc. Welche dieser Informationen bei einer Migration erhalten bleiben müssen und auf welche verzichtet werden kann, sollte im Planungsprozess festgelegt werden (siehe auch Kapitel 3.5.3). Bei der Migration bleibt die Originaldatei in jedem Fall erhalten, d. h. im Hinblick auf den Speicherplatz muss bei jeder Migration mit einer durch das neue Format bedingten Erhöhung des Speichervolumens gerechnet werden.⁹³ Wie hoch diese genau ausfällt, hängt von den Formateigenschaften ab; während sich z. B. bei der Migration von TIFF-Bilddateien in das Format JPEG2000 die Dateigröße verringert, ist bei einer Migration von RAW-Dateien in TIFF mit einer Erhöhung der Dateigröße zu rechnen.

Migrationen finden nicht nur auf Dateiebene statt; auch die zur Langzeitarchivierung verwendete Hardware muss migriert werden, wobei bei der Datenträgermigration nur der Bitstream auf einen neuen Datenträger übertragen wird, z. B. von DVD auf Festplatte.⁹⁴ In diesem Zusammenhang konzentriert sich die wissenschaftliche Diskussion vor allem auf die Wahl des Datenträgers. Da Festplatten nur über eine relativ kurze Lebensdauer von höchstens zehn Jahren verfügen, bleibt auch die Speicherung auf Mikrofilm zumindest was die Lebenszeit des Mediums angeht als Archivierungsstrategie interessant, da dem Mikrofilm eine Haltbarkeit von etwa 500 Jahren zugeschrieben wird; andererseits ist der Zugriff auf die so gespeicherten Daten wiederum mit hohen technischen Anforderungen an (zukünftige) Lesegeräte verbunden.⁹⁵

Bei der Emulation von Daten zum Zwecke der digitalen Langzeitarchivierung wird für die Nutzung der in nicht mehr unterstützten Formaten vorliegenden Daten die nötige Peripherie von Software, Betriebssystem oder Hardware virtuell nachgebaut, um die Daten weiterhin nutzen zu können. Funk führt hierzu als ausführliches Beispiel Emulatoren an, die es ermöglichen, Spiele und Programme des Commodore C-64 auf heute gängigen Betriebssystemen wie Windows und Mac OS ausführen zu können.⁹⁶ Im Gegensatz zur Migration wird immer auf die Originale zugegriffen, d. h. es wird kein zusätzlicher Speicherplatz für die Ablage der migrierten Daten gebraucht und es tritt kein Informationsverlust auf. Für eine Emulation muss also nicht unbedingt die Spezifikation des digitalen Objekts bekannt sein, sondern die der zu emulierenden Umgebung. Ob dies ein Vorteil oder ein Nachteil ist, hängt von der Dokumentation der Originalumgebung, sowie von den technischen Möglichkeiten zur Reproduktion der Umgebung ab. In jedem Fall muss bei jedem größeren Hardware bzw. Betriebssystemwechsel eine neue Emulation programmiert werden⁹⁷, der Aufwand hierfür ist aufgrund der vorgenannten Abhängigkeiten nur schwer plan- und kalkulierbar.

Während bei der Migration die Daten durch die Umwandlung in neue, zum Zeitpunkt der Migration gängige Formate mit Standardsoftware weiterbearbeitet, editiert und ausgetauscht werden können, bleiben die vollständigen Daten bei Emulation nur innerhalb

⁹² Eine aktuelle Liste der von der Library of Congress empfohlenen Dateiformate für die Langzeitarchivierung kann unter <http://www.loc.gov/preservation/resources/rfs/TOC.html> abgerufen werden.

⁹³ Vgl. Funk 2010a, Kap. 8:14

⁹⁴ Vgl. Funk 2010a, Kap. 8:11

⁹⁵ Vgl. Keitel 2010, Kap. 8:32f.

⁹⁶ Vgl. Funk 2010b, Kap. 8:17ff.

⁹⁷ Vgl. Funk 2010b, Kap. 8:23

der Emulationssoftware nutzbar; z. B. könnte bei Emulation einer Umgebung für Word95-Dateien zwar der Text kopiert werden und in eine aktuelle Word-Datei eingefügt werden, stilistische Informationen würden aber nicht übernommen. Noch schwieriger dürfte sich die Übernahme von Bilddateien gestalten – hier ist man bei Emulation unter Umständen tatsächlich auf die Viewer-Ebene beschränkt, eine Übernahme in ein neues Format wäre nur per Screenshot möglich, wodurch originäre Bildinformationen verloren gehen.

Die Erhaltung von (historischer) Hardware und der auf ihr lauffähigen Software wird in der institutionellen Langzeitarchivierung eher als Notlösung praktiziert für die Fälle, in denen derzeit weder Migration noch Emulation möglich ist.⁹⁸

Tendenziell erscheint die Migration unter rationalen Gesichtspunkten besser zu bewältigen als Emulation, da eine Migrationsplanung entsprechend in die Archivplanung und die finanzielle Planung einbezogen werden kann (und muss) und der tatsächliche Aufwand bei der Verwendung gut dokumentierter Formate planbar sein sollte, wohingegen bei der Emulation die organisatorische und finanzielle Planung eine unbekannte Variable darstellt, da Zeitpunkt und Aufwand für die Emulation auf zukünftigen Betriebssystemen vorher nicht genau bestimmt werden können. Jedoch sollte man die beiden Strategien nicht als Gegensätze wahrnehmen: In Notfällen (Migration nicht oder nur mit erheblichem Informationsverlust möglich, Emulation zu aufwändig) kann eventuell immer noch auf den jeweils anderen Ansatz zurückgegriffen werden, da ja in beiden Fällen die Originaldaten vorhanden sind.

3.4.3 Verteilte Speicherung

Unabhängig von der gewählten Archivierungsstrategie muss sichergestellt werden, dass die Daten physisch erhalten bleiben. Dabei besteht der Anspruch, dass die Bewahrung der Daten auch im Fall eines Hardwareausfalls, menschlichem Versagens oder des Eintritts von Elementarschäden, Naturkatastrophen oder Kriegen garantiert ist. Hieraus resultiert die Strategie der verteilten Speicherung der zu archivierenden Daten, d. h. sowohl der Originalobjekte als auch der eventuell migrierten Objekte. In verschiedenen Projekten wurden hierzu unterschiedliche Ansätze erprobt, vom Einsatz eines RAID-Systems, das Daten redundant auf einem virtuellen System bestehend aus verschiedenen physischen Festplatten speichert, bis zur Vernetzung räumlich entfernter Speichersysteme in unterschiedlichen Institutionen, die bei Ausfall eines Systems den Zugriff auf das Objekt über ein vernetztes System ermöglicht.⁹⁹ Letzteres wird im Projekt LOCKSS (Lots Of Copies Keep Stuff Safe) durchgeführt. Die teilnehmenden Institutionen speichern die Daten jeweils in einer sogenannten LOCKSS-Box. Diese tauscht ihre Inhalte mit denen der anderen LOCKSS-Boxen aus, so dass in jeder LOCKSS-Box sowohl die Daten der eigenen Institution, als auch die Daten anderer Institutionen redundant gespeichert sind. Fällt nun die eigene LOCKSS-Box aus, kann automatisch auf die Kopie der benötigten Daten auf einer anderen LOCKSS-Box zugegriffen werden.¹⁰⁰

3.5 Konzepte und Modelle

Die bisherige Langzeitarchivierungsforschung hat verschiedene Konzepte und Modelle hervorgebracht, die auf abstrakter Ebene die derzeit bekannten Probleme der Langzeitarchivierung beschreiben und lösen sollen. Ein immer wieder geäußelter Vorwurf ist

⁹⁸ Vgl. Huth 2010, Kap. 8:24ff.

⁹⁹ Siehe Huth et al. 2010

¹⁰⁰ Vgl. LOCKSS o.J.

jedoch, dass diese Modelle rein theoretisch sind und die Belegung durch praktische Umsetzung fehlt. Einige Modelle sollen im Folgenden vorgestellt werden.

3.5.1 OAIS-Referenzmodell

Das zentrale Modell der digitalen Langzeitarchivierung ist das OAIS-Referenzmodell. OAIS bedeutet Open Archival Information System. Open bedeutet hierbei, dass das Modell nicht hinter verschlossenen Türen, sondern in offenen Foren (weiter-)entwickelt wird.¹⁰¹ Dabei soll es nicht als eine Art Blaupause für mögliche Software-Anwendungen verstanden werden, sondern es beschreibt die Interaktion von Menschen und Systemen, die nötig ist, um digitale Objekte unter Beachtung rechtlicher und organisatorischer Aspekte vertrauensvoll zu archivieren und zugänglich zu machen.¹⁰²

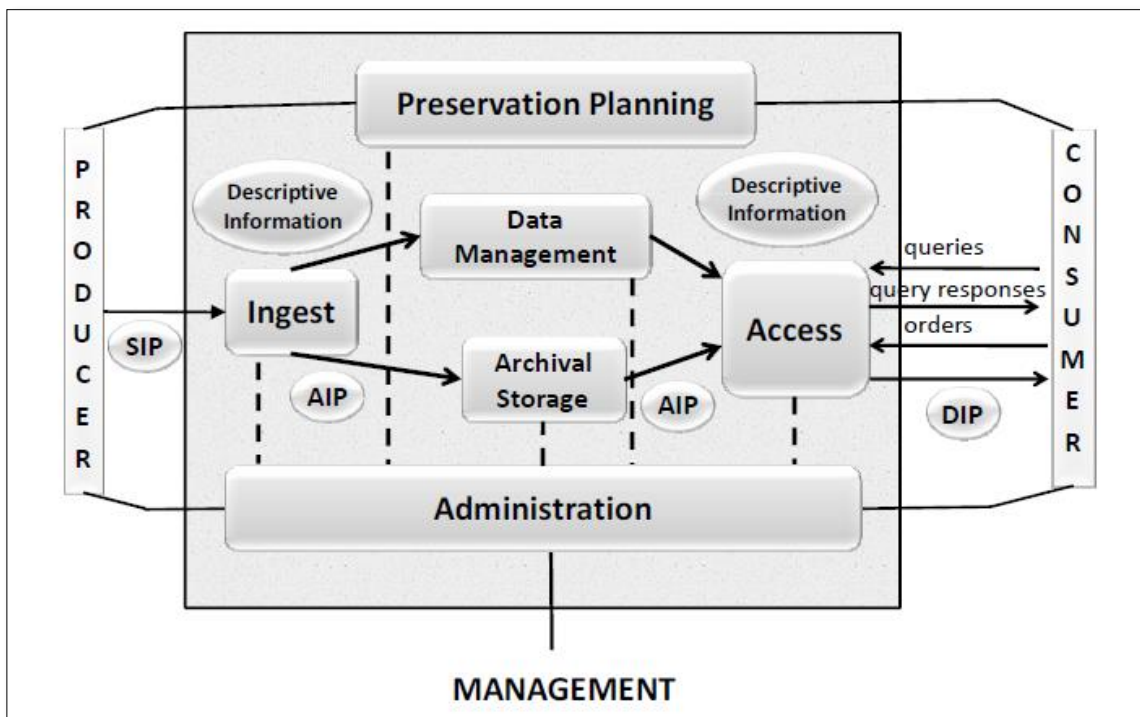


Abbildung 1: Funktionales OAIS-Referenzmodell (Quelle: (CCSDS 2012), S. 4-1)

Das funktionale Modell in Abbildung 1 zeigt die beteiligten Parteien, Informationspakete (IP) und Arbeitsbereiche. Als beteiligte außenstehende Parteien definiert werden die Produzenten der zu archivierenden Objekte (Producer), die Nutzer der archivierten Objekte (Consumer) und das Management, das sich z. B. um externe Angelegenheiten wie Policies, nicht aber um interne Vorgänge kümmert.¹⁰³

Ein Informationspaket besteht aus dem zu archivierenden Objekt und dazugehörigen Metadaten. Die von den Produzenten bereitgestellten Informationspakete werden als Submission Information Package (SIP) bezeichnet. Diese werden dann für die Archivierung aufbereitet, wodurch sie zu Archival Information Packages (AIP) werden. Für den Zugriff durch die Nutzer werden die Daten als Dissemination Information Package (DIP) bereitgestellt.

¹⁰¹ Vgl. CCSDS 2012, S. 1-1

¹⁰² Vgl. Neuroth, Heike et al. 2010; Brübach 2010, Kap. 4:3

¹⁰³ Vgl. CCSDS 2012, S. 2-2f.

Innerhalb des Archivs sind sechs Aufgabenbereiche definiert. Der Bereich *Ingest* umfasst alle für die Übernahme der Objekte in das Archiv nötigen Handlungen, insbesondere die Aufbereitung der SIPs zu AIPs. Hierzu müssen die SIPs erst einmal hinsichtlich ihres Inhalts und Formats überprüft und gegebenenfalls in für die Langzeitarchivierung geeignete Formate migriert werden. Außerdem müssen Metadaten (Descriptive Information) hinzugefügt werden, die das archivierte Objekt hinsichtlich Format, Inhalt und durchgeführter bzw. durchzuführender Erhaltungsmaßnahmen beschreiben.¹⁰⁴ Der Bereich *Data Management* betrifft die Verwaltung der Archivpakete auf Datenbankebene, *Archival Storage* die Speicherung der AIPs und die Überprüfung derselben sowie der verwendeten Datenträger.¹⁰⁵ Übergreifende Bereiche sind *Preservation Planning* und *Administration*. *Preservation Planning* betrifft die Planung der Archivierungsstrategie, eventuell erforderlicher Migrationen und die Überprüfung der technischen Gegebenheiten.¹⁰⁶ Der Bereich *Administration* umfasst die Verwaltung des Archivs an sich sowie der einzelnen Arbeitsschritte von der Aufnahme bis zur Zugänglichmachung der Objekte.¹⁰⁷ Wie in Abbildung 1 durch die gestrichelten Linien dargestellt, hat der Bereich *Administration* Auswirkungen auf alle übrigen Arbeitsbereiche. Der finale Arbeitsbereich ist der Bereich *Access*, bei dem es im Gegensatz zum Bereich *Ingest* um die Zugänglichmachung der archivierten Objekte für die von außen anfragenden Nutzer geht. Dazu gehört die Verzeichnung der archivierten Objekte, damit der Nutzer diese anfragen und abhängig von eventuell gegebenen Zugangsbeschränkungen erhalten kann. In diesem Rahmen werden dann auch die DIPs erstellt.¹⁰⁸

Das OAIS stellt also die komplexen langzeitarchivierungsrelevanten Vorgänge für die institutionelle Archivierung digitaler Objekte in einem Schema dar und bietet ein gemeinsames Vokabular für alle mit der Langzeitarchivierung betrauten Institutionen, wodurch die gemeinsame Kommunikation erleichtert wird.

3.5.2 Designated Community

Um den Wert bzw. die Archivierungswürdigkeit eines digitalen Objekts zu bestimmen, wird in der institutionellen Langzeitarchivierung das Konzept der Designated Community verwendet:

An identified group of potential Consumers who should be able to understand a particular set of information. The Designated Community may be composed of multiple user communities. A Designated Community is defined by the Archive and this definition may change over time.¹⁰⁹

Es handelt sich dabei also um die aktuellen und zukünftigen Rezipienten bzw. Nutzer der archivierten Objekte, z. B. zukünftige Forscher, die auf die Daten und Publikationen der heutigen Zeit zurückgreifen möchten. Die Designated Community kann aber immer nur eine Hypothese sein, da heute nicht mit Sicherheit vorhergesagt werden kann, für welche zukünftigen Nutzergruppen heutige Daten relevant sein werden. Daher schließt die oben genannte Definition auch die Anpassung der durch das Archiv bestimmten Designated Community aufgrund neuerer Erkenntnisse ein. Die nestor-Arbeitsgruppe Policy merkt hierzu an, dass Forschungsbibliotheken wahrscheinlich einfacher ihre De-

¹⁰⁴ Vgl. Neuroth, Heike et al. 2010; Brübach 2010, Kap. 4:10

¹⁰⁵ Vgl. CCSDS 2012, S. 4-2

¹⁰⁶ Vgl. Neuroth, Heike et al. 2010; Brübach 2010, Kap. 4:11

¹⁰⁷ Vgl. CCSDS 2012, S. 4-2

¹⁰⁸ Vgl. CCSDS 2012, S. 4-2f.

¹⁰⁹ CCSDS 2012, S. 1:11

signed Communities bestimmen können als Universitätsbibliotheken, die von Forscherinnen und Forschern verschiedener Fachbereiche in verschiedenem Maße genutzt werden.¹¹⁰ Wichtig ist die Bestimmung der Designated Community nicht nur für die Bewertung der zu archivierenden Objekte, sondern sie spielt auch bei der Bestimmung der Significant Properties eine große Rolle.

3.5.3 Significant Properties

Als Significant Properties werden objektdefinierende Eigenschaften bezeichnet, die für die Rezeption des Objekts durch die Designated Community relevant sind und deshalb unter keinen Umständen im Zuge von Formatmigrationen verloren gehen dürfen.

Das InSpect-Projekt¹¹¹ widmete sich der Untersuchung und Bestimmung dieser Eigenschaften für die Objekttypen Rastergrafiken, E-Mails, Audio-Aufnahmen und strukturierte Texte. Um die Significant Properties eines Objekts zu ermitteln, muss demnach untersucht werden, welche Funktionen das Objekt in seiner vorliegenden Form mitbringt und welche Anforderungen die Designated Community an das Objekt stellt.¹¹² Hierzu werden die technischen Objekteigenschaften ermittelt und kategorisiert. Je nachdem, ob die Objekteigenschaften Informationen zum Inhalt (Content, z. B. Dauer, Umfang), zur Entstehungsumgebung (Context, z. B. Autor, Zeitpunkt), zur Umsetzung (Rendering, z. B. Schrifttyp, Größe), zur Struktur (Structure, z. B. Beziehung zwischen E-Mail und Attachment, Absätze in einer E-Mail) oder zum intendierten Verhalten des Rezipienten (Behaviour, z. B. dass er auf einen angegeben Link in einer E-Mail klickt), enthalten, werden sie entsprechend kategorisiert, um eine bessere Übersicht zu erhalten.¹¹³

Zur Ermittlung der Anforderungen der Designated Community ist es eventuell nötig, dass die ermittelnde Partei tatsächlich einen sogenannten *stakeholder*, also ein Mitglied der Designated Community hinzuzieht.¹¹⁴ Es muss ermittelt werden, welche Handlungen die Designated Community am archivierten Objekt vornehmen möchte, z. B. ob das Objekt nur angezeigt werden soll oder ob auch die Möglichkeit gegeben sein muss, dass das Objekt durch die zukünftigen Nutzer bearbeitet werden kann. Diese Anforderungen an die Nutzbarkeit des Objekts werden dann mit den im vorherigen Schritt festgestellten technischen Objektmetadaten abgeglichen und festgelegt, bis zu welchem Grad die gewünschte Nutzbarkeit bzw. die ermittelten Objektdaten erhalten bleiben müssen.¹¹⁵

Eine Gewichtung der Significant Properties des Objekttyps E-Mail für die Designated Communities „Mitarbeiter der Firma“ und „Finanzverwaltung“ könnte z. B. so aussehen, wobei hier auch die im Rahmen der Fallstudie zum Objekttyp E-Mail¹¹⁶ exemplarisch ermittelten Significant Properties aus der InSpect-Studie mitgenannt werden:

¹¹⁰ Vgl. nestor-Arbeitsgruppe Policy 2014, S. 7

¹¹¹ <http://www.significantproperties.org.uk/>

¹¹² Vgl. Grace et al. 2009, S. 7

¹¹³ Vgl. Grace et al. 2009, S. 9ff.

¹¹⁴ Vgl. Grace et al. 2009, S. 11

¹¹⁵ Vgl. Grace et al. 2009, S. 12f.

¹¹⁶ Vgl. Grace et al. 2009, S. 16ff.

Identifizierte Significant Properties	Significant Properties nach InSPECT	Designated Community	
		Mitarbeiter	Finanzverwaltung
Sender/Empfänger	Local part	10	10
	Domain part	10	10
	Display name	5	5
Betreff der Nachricht	Subject	10	10
Inhaltstext der Nachricht (nur Text)	Message body with no markup	10	10
Kontext einer Nachricht	Keywords	8	4
Datum der Nachricht	Sent date	10	10
Attachments	Attachment	8	10

Tabelle 1: Fallbeispiel Gewichtung von Significant Properties

Hierbei werden z. B. local part und domain part (die beiden Angaben vor und hinter dem @) für beide Communities mit dem Wert 10 angegeben; der angezeigte Absendername (Display name) wird hingegen mit 5 geringer bewertet. Bei einer Formatmigration müssten also die Metadaten „mustermann@abc.de“ unbedingt erhalten bleiben, der Anzeigename „Herr Mustermann“ könnte hingegen wegfallen, z. B. wenn seine Erhaltung als zu aufwändig beurteilt wird.

Insgesamt steht mit dieser Theorie der institutionellen Langzeitarchivierung ein Werkzeug zur Verfügung, das Migrationsprozesse erleichtern kann. Allerdings erfordert es im Vorfeld eine intensive Beschäftigung mit dem zu archivierenden Objekttyp und der Designated Community.

3.6 Technische Unterstützung

Für die institutionelle Langzeitarchivierung existieren bereits Softwarelösungen, die sowohl einzelne Arbeitsschritte wie die Formaterkennung als auch umfangreichere Arbeitsabläufe wie die Archivierungsplanung automatisieren. Diese Lösungen stammen zu einem großen Teil aus den in Kapitel 3.2 erwähnten Projekten und Kooperationen, aber auch kommerzielle Anbieter entwickeln Lösungen für die digitale Langzeitarchivierung. Im Folgenden sollen exemplarisch mit JHOVE eine Lösung für die Formatvalidierung, mit PLATO ein Tool für die Archivierungsplanung und mit Rosetta eine kommerzielle Lösung des Softwareanbieters ExLibris vorgestellt werden.

3.6.1 JHOVE

JHOVE (JSTOR/Harvard Object Validation Environment) ist ein in Zusammenarbeit von JSTOR und Harvard University Library entwickeltes Tool, das für die Formatvalidierung von zu archivierenden Objekten eingesetzt werden kann.¹¹⁷ Außerdem liefert es technische Metadaten und kann so dafür genutzt werden, die als Significant Properties definierten Werte zu überprüfen und zu extrahieren, um diese z. B. in einer externen Beschreibungsdatei des Objekts im Repository zu speichern oder sie für die oben beschriebene Bewertung zu nutzen (vgl. Kapitel 3.5.3). Ähnliche Daten liefern auch DROID und FITS, jeweils in unterschiedlichem Umfang. JHOVE ist eine Java-Applikation und kann sowohl kommandozeilenbasiert ausgeführt werden als auch über eine grafische Oberfläche bedient werden:¹¹⁸

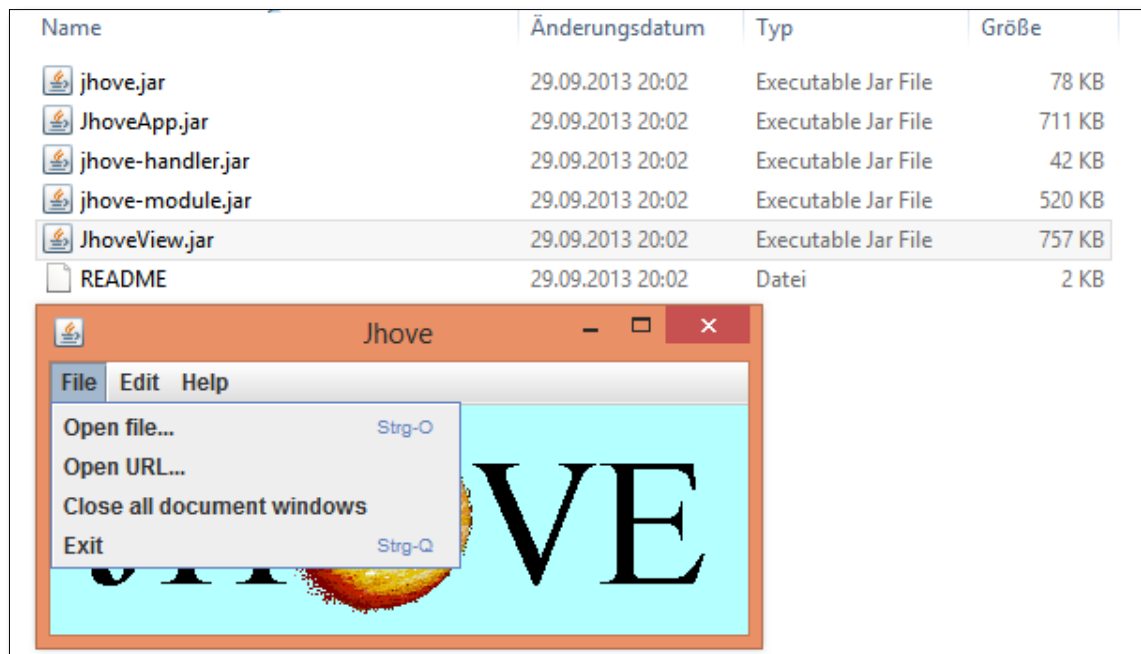


Abbildung 2: Installationsverzeichnis und grafische Oberfläche von JHOVE

Als Webservice kann es in bestehende Workflows bzw. externe Anwendungen integriert werden; es ist z. B. für die schnelle Validierung von Beispieldateien im Planungstool Plato verfügbar, welches im folgenden Kapitel vorgestellt wird.

¹¹⁷ Vgl. JSTOR and the President and Fellows of Harvard College 2012a

¹¹⁸ Vgl. JSTOR and the President and Fellows of Harvard College 2012b

3.6.2 PLATO

Plato ist ein ursprünglich im EU-Projekt PLANETS von Forschern der TU Wien entwickeltes Planungswerkzeug, das seit 2011 im Projekt SCAPE weiterentwickelt wird und aktuell in der Version 4.4 vorliegt. Plato ist als frei zugängliche, auf Java basierende Webapplikation implementiert, die über die Plato-Website¹¹⁹ gestartet werden kann. Ziel des Planungstools ist es, Vorteile und Risiken verschiedener Langzeitarchivierungsstrategien im Vorfeld überprüfen zu können, um eine nachvollziehbare Langzeitarchivierungsplanung zu ermöglichen. Hierzu können sowohl eigene Pläne erstellt als auch vorhandene Beispielpläne und öffentlich freigegebene Pläne anderer Anwender eingesehen werden. Zur Erstellung eines eigenen Plans ist es nötig, dass eine Beschreibung der Kollektion sowie des Planungsvorhabens und die Designated Community angegeben werden, eventuell eine Policy definiert wird sowie mindestens ein Beispielobjekt hochgeladen wird.

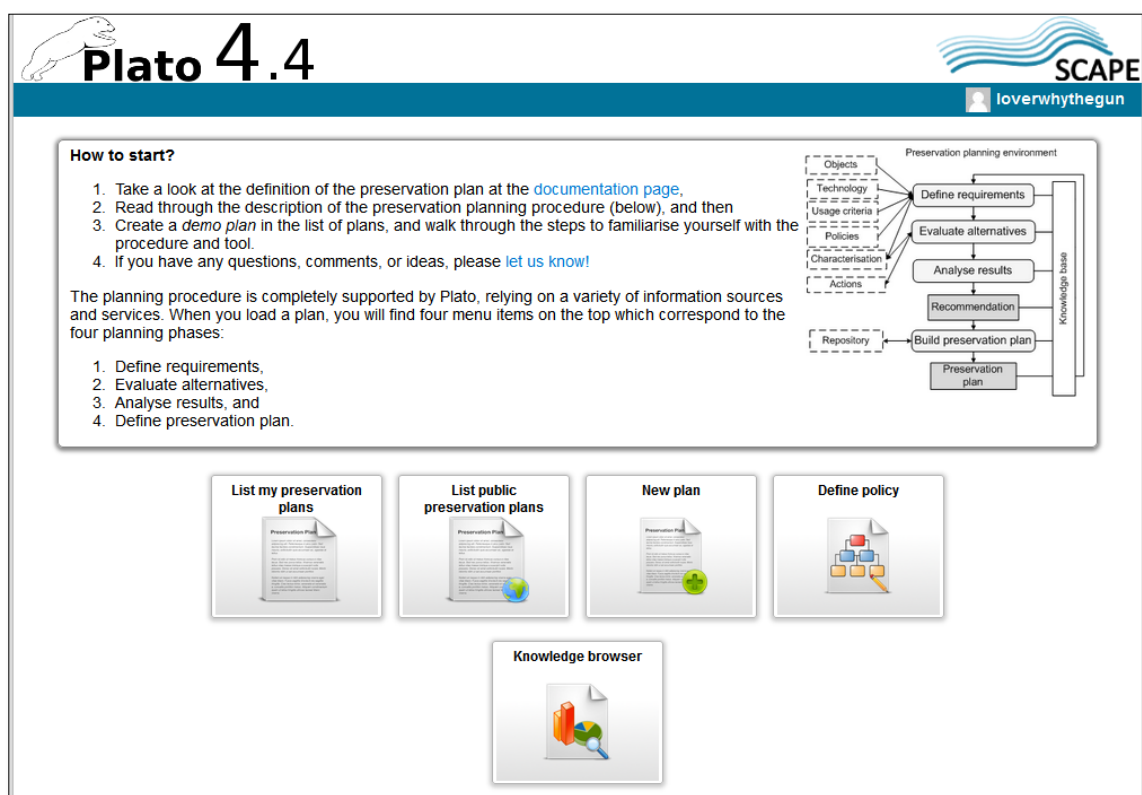


Abbildung 3: Startseite von Plato 4.4

Die einzelnen Planungsschritte *Define requirements*, *Evaluate alternatives*, *Analyse results* und *Define preservation plan* (vgl. Abbildung 3) bauen aufeinander auf, wodurch der Nutzer gezwungen wird, sich an die vorgegebene Abfolge zu halten. Nachdem im Schritt *Define requirements* die oben genannten grundlegenden Angaben sowie die Beispieldateien angegeben wurden, wird der sogenannte *objective tree* erstellt. Dieser besteht im Grunde aus den Significant Properties und kann wie im INSPECT-Projekt vorgeschlagen in die Bereiche Content, Context, Structure und Behaviour aufgeteilt werden (vgl. Kapitel 3.5.3). Durch die per Webservice eingebundenen

¹¹⁹ Siehe TU Wien

Formaterkennungstools DROID, JHOVE und FITS können die technischen Objektmetadaten der Beispieldateien direkt extrahiert werden. Im Planungsschritt *Evaluate alternatives* werden dann die gewünschten Migrationsstrategien ausgewählt und die entsprechenden Migrationsaktionen direkt an den Beispielobjekten erprobt. *Analyse results* erlaubt dann die Analyse der Ergebnisse der Migration. Anhand eines Punktesystems kann leicht abgeschätzt werden, welche Migrationshandlung ein besseres Ergebnis in Bezug auf die im Vorfeld definierten Anforderungen liefert. Die Entscheidung für eine Migrationsstrategie mündet dann in den letzten Schritt *Define preservation plan*.¹²⁰ Der fertige Plan kann als PDF oder als XML-Datei gespeichert und langzeitarchiviert werden.¹²¹

3.6.3 Rosetta

Rosetta ist eine Softwarelösung der Firma ExLibris, die im deutschsprachigen Raum von der Bayerischen Staatsbibliothek, der SLUB Dresden, der ETH-Bibliothek Zürich und dem GoPortis-Verbund zur Langzeitarchivierung eingesetzt wird. Rosetta wurde in Zusammenarbeit mit der National Library of New Zealand entwickelt und 2009 erstmals veröffentlicht, wobei die im OAIS-Modell beschriebenen Funktionen laut Eigenschaft explizit implementiert wurden.¹²²

Tiefergehende Beschreibungen der Systemarchitektur oder Anwenderberichte sind nicht verfügbar, ein bei kommerziellen Produkten häufig anzutreffender Umstand. Laut einer Studie des Fraunhofer-Instituts für Arbeitswirtschaft und Organisation IAO soll aber die Weiterentwicklung durch Anwender z. B. in Bezug auf Schnittstellen möglich sein.¹²³ Hardwaretechnisch können NAS- und Festplattensysteme eingebunden werden.¹²⁴ Für die Langzeitarchivierungsplanung ist eine Schnittstelle zur Formatdatenbank PRONOM verfügbar, über die Informationen zu Formaten und Formatmigrationen abgerufen werden können.¹²⁵ Genauere Informationen zu Implementierungsmöglichkeiten und der Einbindung bestehender lokaler Komponenten sind nicht öffentlich verfügbar.

3.7 Informations- und Schulungsangebote

Wie bereits in Kapitel 3.5.1 erläutert, beschreibt das OAIS nicht eine Software-Architektur, sondern stellt ein Interaktionsmodell von Mensch und Technologie dar. Um entsprechende Handlungen ausführen und Entscheidungen treffen zu können, müssen die mit der Langzeitarchivierung betrauten Personen entsprechend geschult werden.

Deutschsprachige Informationsangebote bieten die Website und die Publikationen von nestor, insbesondere das bereits erwähnte nestor-Handbuch. Ein weiteres zentrales Anliegen des nestor-Kompetenznetzwerks ist die Qualifizierung von Beschäftigten im Bereich Archiv/Bibliothek/Museum für die Aufgaben der digitalen Langzeitarchivierung. Es werden sowohl mehrtägige nestor-Schools als auch eintägige Workshops wie der nestor-Praktikertag angeboten. Ziel von nestor ist es, den erkannten Qualifizierungsbedarf in Kooperation mit den in Deutschland ausbildenden Hochschulen für die Bibliotheks- und Informationswissenschaft abzudecken. Hierzu wurde erstmalig in 2007 ein Memorandum of Understanding unterzeichnet und 2011 erneuert, in dem sich die

¹²⁰ Für eine detaillierte Beschreibung eines Planungsworkflows siehe Engels 2013

¹²¹ Vgl. Kulovits et al. 2010, Kap. 13:4

¹²² Vgl. Peled 2011

¹²³ Vgl. Ferle 2012, S. 118

¹²⁴ Vgl. Ferle 2012, S. 120

¹²⁵ Vgl. Ferle 2012, S. 121

Partner zur kooperativen Entwicklung eines Curriculums sowie zur gegenseitigen Unterstützung bei der Konzeption und Ausrichtung von Schulungen und anderen Qualifizierungsangeboten verpflichten.¹²⁶ So wurden bereits in 2007 e-Learning-Tutorials zu verschiedenen Teilbereichen der institutionellen Langzeitarchivierung von Studenten der Hochschulen Köln, Leipzig, Potsdam und Chur unter didaktischer Mithilfe der Humboldt-Universität Berlin entwickelt, die von den einzelnen Hochschulen innerhalb ihrer Ausbildungsstudiengänge genutzt werden und durch nachfolgende Studienjahrgänge evaluiert und aktualisiert werden.¹²⁷ Der Besuch der nestor-Schools kann z. B. im MALIS-Studiengang der FH Köln als Studienleistung angerechnet werden. Dadurch werden die Erkenntnisse der institutionellen Langzeitarchivierung direkt in die Studiengänge hineingetragen.

Das erst seit 2013 zum nestor-Kompetenznetzwerk gehörende GESIS Leibniz-Institut für die Sozialwissenschaften bietet ebenfalls Qualifizierungsmaßnahmen an. Für interessierte Informationswissenschaftler anderer Institutionen wird pro Jahr ein einführender Workshop veranstaltet, bei dem Basiswissen zu digitalen Langzeitarchivierung vermittelt wird.¹²⁸ Parallel findet auch für Forscherinnen und Forscher, die im Rahmen sozialwissenschaftlicher Forschung empirische Daten erheben, ein Workshop zu Datenmanagement und Archivierung statt.¹²⁹ Das bedeutet, dass hier auch die Weitergabe von Erkenntnissen der institutionellen Langzeitarchivierung an eine Gruppe von Datenerzeugern umgesetzt ist, wie es auch für die Gruppe der privaten Archivierer wünschenswert wäre.

Ein Beispiel für Qualifizierungsansätze auf internationaler bzw. europäischer Ebene ist das durch die europäische Kommission geförderte Projekt DigCurV, in dessen Laufzeit von 2011 bis 2013 sich internationale Partner um die berufliche Qualifizierung in der institutionellen Langzeitarchivierung in Form eines Curriculums bemühten. Ziel ist die Erstellung eines allgemeingültigen Rahmenmodells, nach dem die berufliche Qualifizierung hinsichtlich Langzeitarchivierung vorgenommen werden kann. Dazu wurden drei berufliche Gruppen identifiziert, die in unterschiedlicher Weise an der Langzeitarchivierung beteiligt sind und daher auf unterschiedliche Weise qualifiziert werden müssen.¹³⁰ Eine ähnliche Qualifizierungsstrategie ist auch im nestor-Handbuch beschrieben, woraus dort die oben dargestellte kooperative Qualifizierung abgeleitet wird.¹³¹

3.8 Zusammenfassung

Die institutionelle Langzeitarchivierung verlangt ein streng durchgeplantes, normiertes Vorgehen, dessen Modelle und Konzepte hohe Anforderungen an die durchführenden Einrichtungen und Personen und die verwendeten Systeme stellen. Forschungsprojekte sowie Informations- und Schulungsangebote werden kooperativ organisiert. Die folgende Analyse versucht hieraus Erkenntnisse für die Verbesserung von Personal Digital Archiving abzuleiten.

¹²⁶ Vgl. nestor 2011

¹²⁷ Vgl. nestor 2011, Kap. 19:14f.

¹²⁸ Vgl. GESIS 2014a

¹²⁹ Vgl. GESIS 2014b

¹³⁰ Vgl. DigCurV 2013

¹³¹ Vgl. Scheffel et al. 2010, Kap. 19:6f.

4 VERGLEICHENDE ANALYSE VON PERSONAL DIGITAL ARCHIVING UND INSTITUTIONELLER LANGZEITARCHIVIERUNG

Personal Digital Archiving und institutionelle Langzeitarchivierung beschäftigen sich beide mit dem Erhalt von digitalen Objekten, jedoch unter unterschiedlichen Voraussetzungen und mit unterschiedlichen Anforderungen. Eine grobe Gegenüberstellung der unterschiedlichen Ansätze zeigt die nachfolgende Tabelle:

	Personal Digital Archiving	Institutionelle Langzeitarchivierung
Problemstellung	Archivierung für den jeweiligen Anwender, Zeitraum anwenderbestimmt	Archivierung der sammlungsrelevanten digitalen Daten „für immer“, für bestimmte Gruppe(n) zukünftiger Rezipienten
Wissenschaftliche Diskussion	Einzelne Projekte aus Sicht der institutionellen Langzeitarchivierung, einzelne Studien aus Sicht Personal Information Management/Archiv/Bibliothek	Nationale und internationale Kooperationen, Kompetenznetzwerke und Forschungsprojekte, hoher Professionalisierungsgrad
Sammelrichtlinie	Keine generelle Definition möglich; „alles, was nicht gelöscht wird“	Sammelauftrag der Institution oder spezielle Collection Policy
Wert bzw. Archivierungswürdigkeit	Emotionaler Wert; Wertigkeit für die Selbstdokumentation; veränderlich	Bestimmt durch die Designated Community (zukünftige Forschergenerationen)
Speicherung	HDD, SDD, Cloudspeicher	LOCKSS; RAID-Systeme etc.
Archivierungsstrategien	Bitstream Preservation, kein Bewusstsein für Formate	Bitstream Preservation als Grundlage für Migration oder Emulation
Theorien und Modelle	Zentrales Repository, das Metadaten verwaltet und nötige Archivierungshandlungen meldet und durchführt (Ansatz ähnlich HOPPLA)	OAIS, Designated Community, Significant Properties
Technische Unterstützung	HOPPLA (nicht mehr auf privaten Nutzer ausgerichtet); allgemeine Backup-Lösungen	Tools zur Formaterkennung, Ingest-Tools, PLATO, Rosetta etc.
Informations- und Schulungsangebote	Informationsangebote von NDIIPP, vereinzelt Thementag in öffentlichen Bibliotheken (USA)	Kooperativ, Qualifizierungsmaßnahmen für unterschiedliche Zielgruppen

Tabelle 2: Gegenüberstellung der unterschiedlichen Ansätze

Die dargestellten Gegensätze werden im Folgenden genauer erläutert und eventuelle Anknüpfungspunkte und Handlungsempfehlungen analysiert.

4.1 Problemstellung

Schon in der Problemstellung beider Forschungsbereiche zeichnen sich die unterschiedlichen Voraussetzungen ab: Während für die institutionelle Langzeitarchivierung der Zeitraum der Erhaltung insofern definiert ist, dass eben keine bestimmte Zeitspanne vorgegeben ist, sondern die Objekte auch unter dem Einfluss äußerer Umstände wie Technologiewechsel, Elementarschäden und Kriege möglichst lange erhalten bleiben sollen, kann für Personal Digital Archiving nicht allgemein definiert werden, wie lange die digitalen Objekte erhalten bleiben sollen: Bis sie nicht mehr benötigt werden? Bis zum Tod des Erstellers/Nutzers? Für die nachfolgende Generation? Dieser zeitliche Rahmen ist jeweils individuell. Hinsichtlich der Aufgaben Bewältigung großer Datenmengen, Auswahl geeigneter Datenträger, Formate, Metadaten, der Finanzierung und Organisation ergeben sich hingegen Überschneidungen zwischen beiden Anwendungsbereichen, die aber aufgrund der unterschiedlichen technischen und organisatorischen Möglichkeiten doch wieder einen Gegensatz bilden.

4.2 Wissenschaftliche Diskussion

In der wissenschaftlichen Diskussion von institutioneller Langzeitarchivierung und Personal Digital Archiving zeigt sich die unterschiedliche Gewichtung dieser beiden Forschungsbereiche deutlich: Während sich bei der institutionellen Langzeitarchivierung entsprechende Kompetenznetzwerke herausgebildet haben, die auch länderübergreifend aktiv sind, fehlen in der Diskussion um Personal Digital Archiving solche gemeinschaftlichen Ansätze derzeit noch. Die wissenschaftliche Diskussion wird von Einzelpersonen wie Cathy Marshall und von einzelnen Institutionen wie der Library of Congress/NDIIPP gesteuert. Auch die einzelnen Fachbereiche Bibliothek/Archiv/Personal Information Management sind bisher nicht ausreichend vernetzt, wobei insbesondere die Archivwissenschaft den fehlenden Rückgriff auf Modelle und Verfahren der institutionellen Langzeitarchivierung anprangert, wie das Beispiel Cushing zeigt.¹³² Allerdings gelingt es ihr bisher auch nicht, neue Impulse zu liefern, sondern die bekannten Studien und Projekte bestätigen nur die bereits durch Cathy Marshall publizierten Erkenntnisse. Außer der jährlichen Personal Digital Archiving Konferenz gibt es keine größeren Zusammenkünfte, die auch internationales Publikum ansprechen. Dabei wäre es begrüßenswert, wenn sich die unterschiedlichen Fachrichtungen zusammenschließen würden, um gemeinsam zu erforschen, welche Modelle und Techniken basierend auf dem Archivierungsverhalten der einzelnen Nutzer angewendet werden können, um die Daten jedes Einzelnen zu erhalten.

Eine generelle Beobachtung beim Studium der entsprechenden Fachliteratur zum Thema Personal Digital Archiving ist, dass das Thema oft aus einer sehr persönlichen und emotional geprägten Sicht behandelt wird. Hier könnte ein wenig mehr Abstraktion die wissenschaftliche Beschäftigung mit dem Thema fördern. Andererseits zeigt die persönliche und emotionale Sicht auf das Thema gerade, dass hier jeder betroffen ist, sowohl der professionelle Kurator als auch der Wissenschaftler, der Hobbyfotograf und der Blogger, da alle persönliche Dateien erwerben, erstellen und teilen. Auch Sarah Kim bemerkt in ihrem Überblick zu Aktivitäten und Forschung die sehr persönlich geprägte Herangehensweise an das Thema Personal Digital Archiving und beklagt das Fehlen einer systematischen Herangehensweise.¹³³

¹³² Vgl. Cushing 2010

¹³³ Vgl. Kim 2013, S.276

4.3 Archivierungsverhalten

4.3.1 Archivierungsvorschriften

Die institutionelle Langzeitarchivierung ist aufgrund ihres Anspruchs, transparent, nachvollziehbar und vertrauenswürdig zu sein, sehr stark durch die dokumentierte und automatisierte Auswahl und Verarbeitung der zu archivierenden Objekte bestimmt. Zentrales Dokument ist die Policy, die bestimmt, welche Objekte wie behandelt und archiviert werden. Der Archivierungswert eines Objekts ist in der Regel bereits durch das Sammlungsprofil der Institution bestimmt; durch das Konzept der Designated Community kann dieser Wert nochmals abgestuft und die zu erhaltenden Objektinformationen bzw. Significant Properties entsprechend ermittelt werden.

Auf Personal Digital Archiving kann eine so stringente Archivierungsrichtlinie nicht angewendet werden. Dies würde bedeuten, dass man Individuen vorschreiben müsste, welche Objekte sie zu archivieren haben. Das in den Studien von Marshall, Williams et al. und Copeland beobachtete Archivierungsverhalten zeichnet sich aber gerade dadurch aus, dass den befragten Personen selber nicht klar ist, was sie für die Zukunft archivieren wollen, welchen Wert die gespeicherten Objekte für sie haben und welche Speichermethode die praktikabelste und effizienteste ist.

Vielmehr könnte man hier durch gezielte Information über Speichermethoden, Formate und Wertbemessung dazu beitragen, dass jedes Individuum die für die eigenen Bedürfnisse am besten geeigneten Abläufe ermitteln kann. Zu diesem Ergebnis kommen auch Williams et al.; sie halten insbesondere Hinweise zur Bewertung und Auswahl sowie Lokalisierung und Zugriff auf die digitalen Objekte für hilfreich, aber auch die Ermunterung der Nutzer zur Datenträgermigration bei gleichzeitiger Aufbewahrung der ursprünglichen Trägermedien und ihrer Inhalte.¹³⁴ Auf diesen Ansatz wird in Kapitel 4.6 näher eingegangen.

Copeland regt an, dass speziell zur Vermittlung von Archivierungspraktiken weitere Forschung erfolgen sollte:

Considering the participants' limited knowledge of preservation practices, research is needed on how best to teach individuals to preserve their own digital information. In support of this, future research is needed to understand how public libraries can help contribute to this learning process.¹³⁵

4.3.2 Wert

Wie in Kapitel 2.4.4 erläutert, wird der Wert eines digitalen Objekts bei Personal Digital Archiving nicht wie in der institutionellen Langzeitarchivierung durch abstrakte Größen wie die Designated Community oder einen speziellen Sammelauftrag bestimmt, sondern der Wert eines Objekts wird individuell festgelegt. Dabei sind vor allem emotionale Faktoren, die ebenfalls von außen schwer nachvollziehbar sind, sowie identitätsstiftende bzw. selbstdokumentierende Ansätze Beweggründe für die Speicherung und Erhaltung von digitalen Objekten. Dies macht die Beurteilung für Außenstehende bzw. für anzuwendende Automatismen schwierig. Die in der institutionellen Langzeitarchivierung propagierte Nachvollziehbarkeit ist also auf persönliche Archive kaum anwendbar. Die von Marshall genannten Indikatoren für die persönliche Bewertung von

¹³⁴ Vgl. Williams et al. 2008

¹³⁵ Copeland 2011, S. 1298

digitalen Objekten könnten aber durchaus für manuelle Auswahlprozesse genutzt werden:

- Sortierung der Dateien im Filesystem nach Größe: Selbsterstellte Dokumente mit hoher Dateigröße waren wahrscheinlich aufwändig zu erstellen und haben dadurch Wert;
- Speziell benannte Objekte zwischen Dateien mit Standarddateinamen: Diese müssen einen besonderen Wert (gehabt) haben, da sie durch die Benennung von den anderen Dateien unterschieden und schnell aufgefunden werden können;
- Dateien, die auch außerhalb des Computers verfügbar sind, z. B. in sozialen Netzwerken oder synchronisiert im Cloud-Storage: Diese Dateien müssen durch das Teilen mit anderen bzw. durch die zusätzliche Webspeicherung einen besonderen Wert haben.

Wichtig ist aber nicht nur, in bestehenden Datenpools wertvolle Dateien erkennen zu können, sondern werthaltige Dateien schon beim initialen Speichern mit einem Sonderstatus zu versehen. Diese könnten z. B. direkt an einem besonderen Ort zusätzlich gespeichert werden, wobei dies natürlich auch wieder zusätzlichen Aufwand bedeutet.

Sollten also doch alle Dateien gespeichert und erhalten werden, ohne besondere Auswahl an Hand des Werts? Wie bereits dargestellt, scheint dies derzeit die gängige Methode zu sein. Da sich der Wert nicht immer eindeutig bestimmen lässt und nicht vorhersehbar ist, ob eine Datei irgendwann noch einmal einen besonderen Wert haben wird, scheint dies unter Berücksichtigung des Übermaßes an Speicherplatz keine falsche Methode zu sein. Jedoch stellt sich ja insbesondere das Wiederauffinden von Dateien als schwierig dar, da man nur das Wiederfinden kann, von dem man weiß, dass man es überhaupt hat. Aber wenn z. B. Bilddateien immer nur mit dem Standarddateinamen „*imagexyz*“ benannt sind, ist es sehr aufwändig, die gesuchte Datei tatsächlich zu finden. Die institutionelle Langzeitarchivierung, insbesondere Bibliotheken, kennen dieses Problem schon aus analogen Zeiten. Frühere Zettelkataloge und heutige OPACs und Discovery Services enthalten die deskriptiven Metadaten jedes Objekts im Bestand. So kann nach Autor, Titel, Jahr, Schlagwort etc. gesucht werden und Treffermengen nach Belieben verfeinert werden, bis aus der unbekannten Menge an Objekten das gesuchte Objekt gefunden ist.

Auf Personal Digital Archiving angewandt würde dies voraussetzen, dass auch die digitalen Objekte innerhalb der privaten Sammlung mit entsprechenden Metadaten angereichert werden müssen, was natürlich wiederum Aufwand bedeutet. Dabei existieren heutzutage z. B. für Bilddateien bereits entsprechende Bildbetrachtungs- und Verwaltungstools, die es erlauben, sprechende Dateinamen und rudimentäre Schlagworte in standardisierten Formaten anzulegen.¹³⁶ Durch die Verwendung von Metadaten im Umfeld sozialer Netzwerke, z. B. die sogenannten Hashtags bei Twitter, durch die Tweets als einer bestimmten Diskussion bzw. einem bestimmten Thema zugehörig gekennzeichnet und damit die Auffindbarkeit und Sichtbarkeit des Tweets erhöht wird, kann diesem Ansatz aber eine reelle Chance eingeräumt werden als nützlich und hilfreich erkannt zu werden. In jedem Fall können die Benennung mit sprechenden Dateinamen und die zusätzliche Vergabe von Schlagworten auch die Ergebnisse heutiger Desktopsuchmaschinen verbessern und damit vor dem Verlust von Daten durch Nichtauffindbarkeit schützen.

¹³⁶ Siehe hierzu z. B. Engels 2013 für die Anwendung der Bildverwaltungssoftware Adobe Bridge zur Anreicherung von Metadaten

4.3.3 Speicherorte und Speicherverhalten

Private Anwender sind sich der Gefahr von Datenverlust durch Hardwarefehler durchaus bewusst und reagieren darauf, indem sie Backups durchführen, externe Festplatten einsetzen oder auf Cloud-Speicherdienste wie Dropbox zurückgreifen. Entscheidend bei der Durchführung von Backups ist, dass diese regelmäßig und sorgfältig durchgeführt werden. Wie in Kapitel 2.6.1 beschrieben, verfügen die meisten Betriebssysteme über eine automatische Backup-Funktion. Werden nur individuelle Dateien als Backup auf externen Festplatten gehalten, so muss die Überprüfung und Aktualisierung dieser Dateien manuell stattfinden. In Cloud-Systemen wie Dropbox gespeicherte Dateien synchronisieren sich automatisch. Schwieriger wird es, wenn Dateien nur in sozialen Netzwerken gespeichert sind und keine lokale Kopie besteht, z. B. weil ein direkter Upload vom Smartphone vorgenommen wurde. Diese Dateien werden außerdem oft durch die Service-Anbieter komprimiert, so dass auch beim Rückspeichern der Daten auf die lokale Festplatte nicht das Original zurückgeliefert wird, sondern eine bearbeitete und eventuell durch Qualitätsverlust gekennzeichnete Datei.

Um diese teilweise unfreiwillige verteilte Speicherung besser zu organisieren, kann eventuell tatsächlich die Handhabung innerhalb der institutionellen Langzeitarchivierung zum Vorbild genommen werden, wenn auch in skaliert Form. Wie in Kapitel 3.4.3 erläutert, wird auch in der institutionellen Langzeitarchivierung doppelt und verteilt gespeichert, wie z. B. im LOCKSS-Verfahren. Entscheidend hierbei ist aber, dass es einen zentralen Datenbestand gibt, der bei der besitzenden Institution liegt, d. h. die komplette Kollektion wird immer zusammen gehalten und gespeichert.¹³⁷ Alle vorhandenen Kopien sind Kopien genau dieses Bestands; es gibt keine Kopie bei einer anderen Institution, die Daten enthält, die nicht zum zentralen Bestand gehören. Ähnlich können auch private Anwender verfahren: Zentraler Speicherort ist z. B. die Harddisk des eigenen Computers oder eine externe Festplatte, die im Haushalt aufbewahrt wird und als primärer Datenträger dient. Das bedeutet, dass sämtlich Dateien zuerst hier gespeichert werden, bevor sie auf andere Speichermedien kopiert werden, die dann außerhalb des eigenen Haushalts gelagert werden bzw. der Upload zu Cloud-Speicherdiensten oder in soziale Netzwerke erfolgt. Dateien vom Smartphone sollten zeitnah auch auf diese Festplatte zur primären Speicherung kopiert werden. Für letzteren Fall bestehen bereits entsprechende Synchronisationsoptionen, entweder manuell durch Übertragung der Daten via USB an die Festplatte oder automatisch mittels Cloud-Synchronisierungsdienst, wobei hier natürlich überprüft werden muss, ob eventuell Kompressionsvorgänge stattfinden. Mit Hilfe automatischer Synchronisierungsvorgänge könnte auch Marshalls Einwand begegnet werden, dass ein zentrales Repository aufgrund der verteilten Speicherung nicht realistisch sei.

So kann erreicht werden, dass immer ein zentraler, aktueller Datenbestand zur Verfügung steht, der durch automatische und manuelle Synchronisierung mit den Kopien im Cloud-Storage und auf entfernt gelagerten externen Festplatten jederzeit wiederhergestellt werden kann. Außerdem kann so immer auf die unkomprimierten Originale zugegriffen werden sowie Datenverlust durch Einstellung eines webbasierten Dienstes vermieden werden.

¹³⁷ Vgl. Cushing 2010, S. 306

4.3.4 Speichern/Langzeitarchivieren

Ein Aspekt, der sowohl von den Teilnehmern der jeweiligen Studien als auch von den untersuchenden Wissenschaftlern wenig beachtet wird, ist die fehlende Trennschärfe zwischen den Begriffen Speichern (*to save/to store*) und Archivieren (*retention/preservation*). Speichern könnte man im Rahmen des Personal Digital Archiving vielleicht besser mit „nicht löschen“ umschreiben: Jede digitale Datei muss in irgendeiner Form gespeichert werden, sei es passiv als automatisches Zwischenspeichern durch das zur Erstellung genutzte Programm oder aktiv durch den Ersteller im Rahmen verschiedener Versionen einer Datei innerhalb des Erstellungsprozesses. Oft werden Dateien danach nicht gelöscht, aus den schon benannten Gründen wie hohe Verfügbarkeit von Speicherplatz, aber auch weil nach dem letztmaligen Speichern und der anschließenden Verwendung (z. B. Upload in ein soziales Netzwerk, Versand per Mail, Ausdruck) der Aufwand, zur Datei im Filesystem zurückzugehen und diese explizit zu löschen unnötig hoch erscheint. Während bei analogen Formen wie z. B. Ausdrucke oder Notizzetteln die Entwurfsfassungen in den Papierkorb geworfen werden, sobald diese überholt sind und sich auf dem Schreibtisch stapeln, tritt dieser haptische Effekt bei digitalen Dokumenten nicht auf.

Das beschriebene Verhalten macht es schwierig, zwischen den Objekten zu unterscheiden, die tatsächlich in den Augen des Nutzers einen subjektiven Langzeitarchivierungswert haben und jenen, die nur aus Bequemlichkeit gespeichert wurden.

Auch ist unklar, ob der einzelne Anwender sich unter dem Konzept „Langzeitarchivierung“ wirklich etwas vorstellen kann und seine Dateien als derart wichtig wahrnimmt, um überhaupt den Gedanken zu fassen, diese für einen längeren Zeitraum behalten zu wollen. Oder der Anwender geht davon aus, dass seine Dateien ihm sowohl im Hier und Jetzt als auch in Zukunft immer zur Verfügung stehen werden, da er Probleme wie Formatwandel, fehlende Softwareunterstützung etc. nicht kennt bzw. nicht wahrnimmt.

In jedem Fall kann auch hier nur die vorerst praktikabelste Maßnahme sein, dass Anwender entsprechend aufgeklärt werden, so dass sie selber in der Lage sind, ihr Archivierungsverhalten auf ihre Objekte und Anforderungen abzustimmen.

4.3.5 Daten- und Formatmigration

Die in Kapitel 2.4.3 vorgestellten Studien von Marshall zeigen, dass private Anwender im Rahmen der Datenmigration in der Regel die vorhandenen, nicht bewerteten Dateien bei Wechsel des PC ohne weitere Untersuchung der Daten auf den neuen Rechner kopieren, aber auch Daten explizit auf alten Computern belassen, um bei Bedarf darauf zugreifen zu können. Letzteres hat auch den Hintergrund, dass Nutzer bereits die Erfahrung gemacht haben, dass Dateien auf neuen Computern mit neuer Software nicht mehr geöffnet werden können und hier schließlich die Dateien und die benötigte Software zusammen bereitstehen, solange der Computer noch funktioniert.¹³⁸ Das bedeutet, dass der private Anwender an Stelle von Migration oder Emulation von Formaten das Computermuseum als für ihn einfachste Strategie für den Zugriff auf ältere Dateien wahrnimmt. Somit stellt also die Ausnahme in Sachen Erhaltungsstrategien innerhalb der institutionellen Langzeitarchivierung im Personal Digital Archiving Umfeld die Regel dar. Auch alte Datenträger werden nicht gegen neuere getauscht, sondern so lange mitsamt der Daten aufbewahrt, wie sie angesprochen werden können.

¹³⁸ Vgl. Marshall 2008a

Ob hier tatsächlich ein Umdenken angestoßen werden kann, ist fraglich. Die in der institutionellen Langzeitarchivierung angewendeten Strategien Migration und Emulation setzen einen hohen Aufwand voraus, z. B. Bestimmung der Significant Properties für die Migration oder Erstellung/Nutzung von Emulationssoftware. Wie bereits gezeigt, ist aber gerade Aufwand ein Hindernis für die Durchführung von Personal Digital Archiving Praktiken. Von daher scheint der Ansatz von Marshall, Dateien regelmäßig auf der Metadatenebene zu überprüfen und Migrationen automatisch vorzunehmen, wie er auch in HOPPLA umgesetzt ist, generell der Richtige zu sein, da hierbei für den Anwender kein unnötig hoher Aufwand entsteht. Solange dies nicht realisierbar ist, sollten Nutzer animiert werden, ihre Daten in offenen Standardformaten zu speichern, um zu gegebener Zeit eine Migration bzw. Emulation zu ermöglichen.

4.4 Konzepte und Modelle

Hinsichtlich etablierter Konzepte und Modelle zeigen sich extreme Unterschiede zwischen institutioneller Langzeitarchivierung und Personal Digital Archiving. Während die institutionelle Langzeitarchivierung mit dem OAIS-Referenzmodell über einen zentralen theoretischen Standard verfügt, der der Ausarbeitung von Workflows und der Entwicklung entsprechender Softwarelösungen zugrunde liegt und dessen Vokabular sich als gemeinsame Sprache innerhalb der Langzeitarchivierung etabliert hat, verfügt Personal Digital Archiving nicht über ein zentrales Modell. Hierfür können zwei verschiedene Ursachen angenommen werden: Zum einen gibt es bisher, wie ja bereits von Sarah Kim bemängelt, zu wenig echte Grundlagenforschung zum Thema Personal Digital Archiving, als dass sich daraus bereits übergreifende Theorien und Modelle ableiten ließen. Zum anderen zeigt die bisher getätigte Forschung keine echte gemeinsame Grundlage der Archivierungsaktivitäten von Privatpersonen, von der ein allgemeingültiges theoretisches Modell abgeleitet werden könnte.

Die eigentliche Frage, die es zu beantworten gilt, ist aber eher: Braucht Personal Digital Archiving wirklich eine für Außenstehende so monströs erscheinende theoretische Grundlage wie das OAIS-Referenzmodell? Das OAIS-Referenzmodell beschreibt einen sehr komplexen Vorgang, der der zu erwartenden Datenmenge, Ansprüchen an die Vertrauenswürdigkeit und Transparenz sowie der Organisation von Geschäftsgängen geschuldet ist. Natürlich sollten diese Gesichtspunkte auch in der privaten Archivierung beachtet werden, jedoch in angepasster Form. In erster Linie erscheint es wichtig, dass Thema Personal Digital Archiving überhaupt in die öffentliche Diskussion zu bringen. Einfache Hilfsmittel und Anleitung können hier wahrscheinlich mehr bewirken als eher abschreckende theoretische Modelle.

4.5 Technische Unterstützung

Hinsichtlich der technischen Unterstützung der Archivierungsbedürfnisse privater Anwender stellt sich wiederum das Archivierungsverhalten als Problem dar. Solange Dateien verteilt und ohne zentrale Datenbasis gespeichert werden, kann eine Software wie HOPPLA nicht auf den kompletten Datenbestand angewendet werden. Der von Cathy Marshall beschriebene und mit HOPPLA auch grundsätzlich umgesetzte Ansatz, Dateien auf der Metadatenebene zu überprüfen, bezieht aber auch solche Objekte mit ein, die sich an Webspeicherorten befinden. Hierzu müsste eine Software wie HOPPLA Zugriff auf die dort gespeicherten Daten haben, was eine entsprechende Programmschnittstelle und die Übergabe persönlicher Zugangsdaten an die Software voraussetzen würde.

Dass HOPPLA derzeit generell nicht weiter entwickelt wird und der Fokus zuletzt auch eher auf kleinere Unternehmen als auf den Privatanwender gerichtet wurde, ist

bedauerlich. Allerdings: Damit eine Software wie HOPPLA von Privatanwendern eingesetzt werden kann, muss auch hier erst einmal das Problembewusstsein vorhanden sein.

Generell zeigt sich sowohl in der institutionellen Langzeitarchivierung als auch im Personal Digital Archiving, dass für einzelne Aufgaben spezielle Software vorhanden ist oder bereits vorhandene Software für bestimmte Aufgaben genutzt werden kann. Systeme, die übergreifend Aufgaben bearbeiten können, existieren als Forschungsprojektoutput (HOPPLA) bzw. als kommerzielle Lösung für die institutionelle Langzeitarchivierung (Rosetta), die bereits von mehreren Forschungsbibliotheken im deutschsprachigen Raum eingesetzt wird, wobei durch die wenig detaillierte, eher marketingorientierte Dokumentation nicht geklärt werden kann, wie das System genau funktioniert. Da für die institutionelle Langzeitarchivierung neben Rosetta kaum übergreifende Lösungen bestehen und viele Institutionen bestrebt sind, statt der Einführung einer neuen Lösung lieber bestehende Lösungen als modulare Komponenten eines eigenen Systems einsetzen möchte, entwickeln sich auch hier kooperative Lösungen wie z. B. die vom Historischen Archiv Köln in Zusammenarbeit mit der Stadt Köln und anderen regionalen Partnern entwickelte Ingest-Lösung DiPS, die auf vorhandenen Lösungen von HP/SER basiert.¹³⁹

Beide Bereiche stehen also vor demselben Dilemma: Aufgrund der Datenmenge lassen sich viele Vorgänge nur automatisch bearbeiten, wobei vorhandene Lösungen nur Teilbereiche abdecken können. Der institutionellen Langzeitarchivierung kommt hierbei zugute, dass mit dem OAIS-Referenzmodell ein Standard vorliegt, der die benötigten Komponenten und Abläufe benennt und definiert. Für die private Archivierung liegen bisher keine brauchbaren Modelle vor, an der sich eine mögliche Softwareentwicklung orientieren kann. Eine Lehre kann aber aus der institutionellen Langzeitarchivierung gezogen und auch auf Personal Digital Archiving angewendet werden: Lösungen sind in der Regel nur im Verbund möglich, da sie Fachwissen aus verschiedenen Bereichen (Archiv/Bibliothek/IT) erfordern und hohen personellen wie finanziellen Aufwand erfordern. Dies könnte erklären, wieso bisher keine kommerziellen Lösungen für den Privatanwender bereitstehen. Und selbst mit den entwickelten Lösungen bleiben viele Aufgaben bestehen, bei denen manuelle Eingriffe des Anwenders nötig sind, d. h. eine Softwarelösung wird das Problem Personal Digital Archiving sicher auch nicht vollautomatisch lösen. Daher ist auch hinsichtlich der technischen Aspekte eine Qualifizierung des Anwenders nötig.

4.6 Informations- und Schulungsangebote

Die bisherigen Informations- und Schulungsangebote zu Personal Digital Archiving fallen trotz des in den vorgestellten Studien angemerkten Bedarfs relativ dürftig aus. Derzeit sind nur englischsprachige Informationen vorhanden, wobei sich nur die Library of Congress/NDIIPP explizit an den privaten Nutzer wendet und die entsprechenden Informationen leicht zugänglich auf ihrer Website bereithält. Die anderen englischsprachigen Angebote sind Ergebnisse aus Projekten, die vor allem aus der Perspektive der späteren Übernahme privater Daten in institutionelle Repositorien durchgeführt wurden. Sie sind daher zum einen schwerer recherchierbar und zum anderen nicht unbedingt auf die Fähigkeiten privater Anwender ausgelegt. Generell fehlen gemeinschaftliche Ansätze, wie es sie in der institutionellen Langzeitarchivierung im nestor-Kompetenznetzwerk gibt. Wie auch bei der Bewältigung der digitalen Langzeitarchivie-

¹³⁹ Vgl. Huppertz 2013

rung an sich bieten sich auch bei der Qualifizierung kooperative Ansätze an. Bezogen auf Personal Digital Archiving wäre hier vielleicht nicht die Kooperation einzelner Institutionen in der Herstellung der Informationen gefragt, sondern zentral erstellte Informationen müssten durch entsprechende Institutionen an potenzielle Nutzer herangetragen werden. In Deutschland könnten zum Beispiel von nestor entsprechende Materialien in Form von Flyern und einer Website bereitgestellt werden. Diese Informationen könnten dann sowohl von wissenschaftlichen Bibliotheken als auch von öffentlichen Bibliotheken für die Gestaltung eigener Angebote genutzt werden.

Zusätzlich muss auch der Aspekt der Qualifizierung der vermittelnden Bibliothekare beachtet werden.

4.7 Fazit der Analyse

Die Analyse zeigt, dass derzeit die Anknüpfungspunkte zwischen dem durch relativ heterogenes Nutzerverhalten gekennzeichnetem Personal Digital Archiving und der streng reglementierten institutionellen Langzeitarchivierung relativ gering sind. Dies ist insbesondere bedingt durch die eingeschränkten technischen und fachlichen Möglichkeiten und Kompetenzen privater Anwender. Abhilfe geschaffen werden könnte durch eine bessere Aufklärung der privaten Anwender über Langzeitarchivierungsmaßnahmen. Entsprechende Hinweise könnten durch nestor erstellt und durch wissenschaftliche und öffentliche Bibliotheken weitergegeben werden. Diese Hinweise sollten enthalten:

- Informationen zur Auswahl/Werterkennung:
 - Große Dateien
 - Spezielle Benennung
 - Geteilte Dateien
- Informationen zur Identifikation der Daten/Verbesserung der Desktopsuche:
 - Sprechende Dateinamen
 - Schlagworte, Tags
- Informationen zu Formaten (offene Standards)
- Informationen zu Speichermethoden (primärer Speicherort, zusätzliche Kopie außerhalb der Wohnung, Cloudspeicher)
- Informationen zu Hardwaremigration

Diese Informationen stellen primäre Handlungsoptionen dar, d. h. diese Maßnahmen können direkt auf die gesammelten Daten angewendet werden. Natürlich könnten hier noch weitere Punkte aufgezählt werden, wie die Festlegung einer privaten Policy, Anlage von Index-Dokumenten etc. Da diese aber eher als sekundäre Maßnahmen einzustufen sind – es wird nicht direkt mit den Objekten gearbeitet, sondern es werden Metadokumente erstellt – erfolgt eine Beschränkung auf die primären Maßnahmen. Die mögliche zukünftige Erstellung und Verwendung dieser Informationen wird im folgenden Kapitel näher untersucht.

5 VORSCHLÄGE FÜR DIE ZUKÜNFTIGE VERMITTLUNG VON PERSONAL DIGITAL ARCHIVING IN DEUTSCHLAND

Als Fazit des Vergleichs von Personal Digital Archiving und institutioneller Langzeitar Archivierung wurde die Notwendigkeit von Informations- und Schulungsangeboten zu Personal Digital Archiving erkannt, da entsprechende technische Lösungen fehlen und die nötige Skalierbarkeit der Methoden für eine hinsichtlich technischem und fachlichem Verständnis heterogene Zielgruppe nicht durch einen allgemeinen Ansatz erreicht werden kann. Wie in Kapitel 2.7 erläutert, fehlen derzeit in Deutschland entsprechende Informations- und Schulungsangebote zum Thema Personal Digital Archiving. In diesem Kapitel sollen drei Vorschläge für zukünftige Informations- und Schulungsmöglichkeiten für verschiedene Zielgruppen aufgezeigt werden: Allgemeine Informationen für die breite Öffentlichkeit in Form von Broschüren und Web-Angeboten, Informationskompetenzschulungen für Nutzer wissenschaftlicher Bibliotheken und Angebote für die Nutzer öffentlicher Bibliotheken.

5.1 Personal-Digital-Archiving-Hinweise für die breite Öffentlichkeit

Derzeit existieren keine deutschsprachigen Informationen zu Personal Digital Archiving von offizieller Seite, z. B. von nestor oder der Deutschen Nationalbibliothek. In Form von Flyern, Broschüren oder einer Website könnten entsprechende Informationen leicht zugänglich bereitgestellt werden. Hinsichtlich der Inhalte könnte auf bereits bestehende englischsprachige Angebote zurückgegriffen werden.

Wie bereits im Kapitel zu Schulungsaktivitäten und in der Analyse erläutert, wurden in verschiedenen Projekten und Initiativen Richtlinien und Hinweise für Privatanwender zur Archivierung ihrer digitalen Sammlungen erstellt: Im InterPARES2-Projekt¹⁴⁰ (2007), im PARADIGM-Projekt¹⁴¹ (2008) und durch die NDIIPP/Library of Congress¹⁴² (2010). Diese Richtlinien könnten als Grundlage für entsprechende deutschsprachige Informationen dienen, die interessierten Anwendern zur Verfügung gestellt werden könnten. Die drei Publikationen weisen allerdings Unterschiede in Aufmachung, Detailliertheit und Anforderung an den privaten Anwender auf. Aufgrund der in der Darstellung und Analyse gewonnenen Erkenntnissen zu Personal Digital Archiving und den Ansprüchen an Nutzerfreundlichkeit, Skalierbarkeit und Verständlichkeit sollen die einzelnen Hilfsmittel im Folgenden bewertet und ihre deutschsprachige Umsetzbarkeit erörtert werden. Dabei wird insbesondere überprüft, ob die im vorausgehenden Kapitel festgelegten Inhalte in den jeweiligen Informationsangeboten bereits berücksichtigt sind.

¹⁴⁰ Siehe InterPARES 2 Project 2007

¹⁴¹ Siehe PARADIGM 2008a

¹⁴² Siehe NDIIPP 2010

5.1.1 InterPARES2 Creator Guidelines: Making and maintaining digital materials: Guidelines for individuals

Trotz ihres in der Überschrift genannten Anspruchs „guidelines for individuals“ weisen diese Richtlinien sehr viel Anteil an Fließtext sowie eine große Detailliertheit auf. Nach einer Einleitung und der Definition einiger Fachbegriffe (record, information, authenticity etc.) folgen zehn nummerierte Abschnitte, ein Fazit und eine Kurzvorstellung des InterPARES-Projekts. In den einzelnen Abschnitten wird ausführlich auf Probleme wie die Auswahl der Hardware/Software/Formate¹⁴³, Identifikation der Daten durch deskriptive Metadaten¹⁴⁴ und Sicherung der Daten gegen unautorisierten Zugriff¹⁴⁵ eingegangen; aber es werden auch Konzepte wie die Beständigkeit¹⁴⁶ von Daten erläutert. Insgesamt wird sehr viel Information bereitgestellt, allerdings ist unklar, wer die Zielgruppe dieser Handreichungen sein soll. Im Text selber heißt es:

These guidelines have been developed for individuals who create digital materials in the course of their professional and personal activities to help them make informed decisions about making and maintaining these materials in ways that will help ensure their preservation for as long as they are needed. They may also be useful for small organizations or groups of individuals, such as medical offices, consulting groups, or teams of research scientists.¹⁴⁷

Entgegen der obigen Aussage scheint aber eher der professionelle Anwender im Vordergrund zu stehen, da das vermittelte Wissen sehr umfangreich und die Tipps für einen privaten Anwender sehr aufwändig umzusetzen sind. Zum Beispiel wird im Abschnitt zu Software, Hardware und Formaten geraten, Änderungen an der Software im Quellcode zu dokumentieren – ein normaler Privatanwender wird wohl keine Änderungen vornehmen und könnte dies auch nicht im Quellcode dokumentieren – aber auch die vorgeschlagene Dokumentation des kompletten Systems würde für einen Privatanwender einen sehr hohen Aufwand bedeuten, dessen Nutzen unklar ist:

These basic specifications will ensure that those who come after you understand the context in which you are working now.¹⁴⁸

Es wird also davon ausgegangen, dass private Archivierung nicht nur zum Selbstzweck stattfindet, sondern es nachfolgende Generationen gibt, die sich für die Daten interessieren. Da sich das InterPARES-Projekt auch vornehmlich mit der Archivierung durch Privatpersonen vor dem Hintergrund der späteren Übernahme durch eine Gedächtnisinstitution beschäftigt hat, liegt der Verdacht nahe, dass die Anforderungen aus Sicht der Institutionen formuliert wurden, ohne besondere Skalierung hinsichtlich der Fähigkeiten des jeweiligen Produzenten.

Hinsichtlich der zu vermittelnden primären Maßnahmen zeigen sich die Guidelines uneinheitlich: Während Speichermethoden und Hardwareupdates zwar erläutert, aber keine konkreten Tipps gegeben werden, sind die Hinweise zur Dateiidentifikation zu ausführlich. Einzig bei den Hinweisen zu offenen Formaten werden auch konkrete Tipps gegeben. Informationen zu Auswahl und Bewertung sind nicht enthalten.

¹⁴³ Vgl. InterPARES 2 Project 2007, S. 3

¹⁴⁴ Vgl. InterPARES 2 Project 2007, S. 5

¹⁴⁵ Vgl. InterPARES 2 Project 2007, S. 7

¹⁴⁶ Vgl. InterPARES 2 Project 2007, S. 4

¹⁴⁷ InterPARES 2 Project 2007, S. 2

¹⁴⁸ InterPARES 2 Project 2007, S. 3

5.1.2 PARADIGM Guidelines for creators of personal archives

Die PARADIGM Guidelines for creators of personal archives starten mit einer Einleitung in Form von praktischen Tipps zur Anwendung der folgenden Hinweise und wenden sich dann den Punkten Organisation und Dateibenennung, E-Mail-Management, Formate und Software, Backup, Hardwarepflege, Schutz des Systems, Passworte und Verschlüsselung, Eigentumsrechte, Technologiewechsel zu, nicht ohne zum Schluss noch den Hinweis zu geben, dass digitale Archivare, also Archivare, die auf die Bewahrung digitaler Daten spezialisiert sind, ebenfalls als Informationsquelle zur Verfügung stehen. Die einzelnen Abschnitte bestehen jeweils aus Fließtext und konkreten Anweisungen in Aufzählungsform. Inhaltlich werden auch hier viele zusätzliche Details erläutert, wie z. B. Vorteile von Open Source Software und klare Formatempfehlungen ausgesprochen, z. B. für den offenen Standard ODF für Textdokumente, MySQL für Datenbanken und TIFF für Rasterbilddateien. Allerdings wird in der Einleitung explizit darauf hingewiesen, dass nicht alle Punkte umgesetzt werden müssen, sondern der jeweilige Anwender diese Tipps entsprechend seines Umfelds und seiner Fähigkeiten benutzen kann:

It is important to choose solutions which suit you, so adopt – and adapt – these according to your needs.¹⁴⁹

Außerdem wird noch eine übersichtliche Tabelle bereitgestellt, die die elf wichtigsten Tipps zur Bewahrung persönlicher Daten enthält und zu jeder durchzuführenden Handlung die wichtigsten Informationen und eine Angabe zur Frequenz der Ausführung der Handlung (ständig, wöchentlich, etc.) enthält.¹⁵⁰

Neben den Hinweisen zu Formaten enthalten die Guidelines auch zu allen anderen definierten primären Maßnahmen Informationen. Hinsichtlich Speichermethoden wird die zusätzliche Speicherung auf externen Medien, an entfernten Orten und in der Cloud empfohlen. Zur Dateidentifikation werden sprechende Dateinamen, eine definierte Ordnerstruktur und Metadaten empfohlen. Hardwareupdates werden ebenfalls erwähnt. Hinsichtlich Auswahl/Bewertung wird nur der schlichte Hinweis „delete what’s not important“¹⁵¹ geliefert.

Insgesamt setzen diese Tipps und Hinweise die analysierten Anforderungen ansprechend und für die Zielgruppe Privatanwender geeignet um.

5.1.3 NDIIPP Preserving your digital memories

Ein ganz anderer Aufbau zeigt sich bei den Archivierungshinweisen der Library of Congress/ NDIIPP. Nach einem kurzen erläuternden Absatz zur Bewahrung digitaler Informationen und dem Satz „But preserving digital information is a new concept that most people have little experience with“¹⁵² als Begründung für die Bereitstellung der Hinweise werden jeweils auf einer Seite Hinweise zur Archivierung der digitalen Objekttypen Fotos, Audio, Video, E-Mail, persönliche Dokumente und Websites gegeben. Diese Hinweise sind auch als Einzeldokumente auf der Website abrufbar. Die Hinweise zu jedem Objekttyp beginnen mit einer kurzen Einleitung, gefolgt von den jeweiligen Tipps in Spiegelstrichen, sortiert nach der organisatorischen Reihenfolge der einzelnen

¹⁴⁹ PARADIGM 2008a

¹⁵⁰ Vgl. PARADIGM 2008b

¹⁵¹ PARADIGM 2008a

¹⁵² NDIIPP 2010, S. 2

Arbeitsschritte. Obwohl es um verschiedene Objekttypen geht, folgen die Hinweise in der Regel einem grundlegenden Aufbau:

- Identify: Der Anwender soll in diesem Schritt identifizieren, wo sich z. B. seine digitalen Fotos befinden (Gerät, Computer, externe Speichermedien, Web).
- Decide: Der Anwender soll entscheiden, welche Dateien er für wichtig hält und diese auswählen, egal ob es sich um einige wenige oder um viele Dateien handelt. Je nach Objekttyp wird auch empfohlen, bei mehreren Kopien die mit der höchsten Qualität zu wählen (Fotos, Audio, Video) oder finale Versionen und Drafts (Video, persönliche Dokumente).
- Organize: Der Anwender soll den Dateien beschreibende Dateinamen geben und sie mit den Namen von gezeigten Personen und beschreibenden Schlagwörtern versehen (Foto, Video) bzw. weiteren zusätzlichen Informationen (Audio). Außerdem wird empfohlen, eine Verzeichnis bzw. eine Ordnerstruktur auf dem Computer anzulegen und diese und die abgelegten Objekte kurz zu beschreiben.
- Make copies: In diesem Abschnitt wird dem Anwender empfohlen, mindestens zwei Kopien zu machen, davon eine auf dem Computer zu halten und die andere auf einem externen Speichermedium oder online zu speichern. Die externen Speichermedien sollen an einem anderen Ort aufbewahrt werden, „as physically far apart as practical.“¹⁵³ Außerdem soll die Beschreibung der Verzeichnisstruktur und Objekte mit anderen wichtigen Papieren an einem sicheren Ort aufbewahrt werden. Die gespeicherten Dateien sollen einmal jährlich auf Lesbarkeit überprüft werden. Es wird weiter empfohlen, alle fünf Jahre bzw. wenn nötig neue Kopien anzulegen um Datenverlust zu vermeiden.

Für Audio, E-Mail, Websites und Dateien in sozialen Netzwerken gibt es noch den zusätzlichen Schritt Export, der Hinweise zur lokalen Speicherung der Dateien enthält.

Diese Hinweise sind jeweils als einfache Anweisungen in einem Satz formuliert („Pick the images“, „Create a directory“) wobei durchgängig Imperative verwendet werden; Handlungsoptionen werden nur bei der Menge der auszuwählenden Dateien und der Speicherung auf dem eigenen PC gegeben („You can pick a few photos or many“, „One copy can stay on your computer or laptop“). Durch diese klaren Anweisungen werden keine verwirrenden Handlungsoptionen zur Auswahl gestellt, sondern der Anwender findet auf einer Seite klare Arbeitsschritte und Abläufe vor.

Inhaltlich fällt auf, dass keine Fremdworte wie z. B. „metadata“ verwendet werden, sondern dieser Begriff mit „descriptive subjects“ umschrieben wird. Formate und Formatmigrationen werden nicht empfohlen, die Archivierungsstrategie stützt sich auf Bitstream Preservation und redundante Speicherung.

Hinsichtlich der im Analyseteil gewonnenen Erkenntnisse stellen diese Hinweise die einfachste und am leichtesten umsetzbare Variante dar. Der Anwender kann praktisch mit einer der Hinweisseiten zur Hand direkt mit der Archivierung beginnen; ein größerer theoretischer Hintergrund oder technische Erläuterungen werden nicht vermittelt, wodurch die einzelnen Handlungsschritte in den Vordergrund rücken. Der objektorientierte Ansatz erhöht die Übersichtlichkeit und ist wahrscheinlich auch für die meisten Anwender als Einstieg besser geeignet, da sie z. B. mit dem Vorsatz ans Werk gehen, jetzt endlich einmal ihre digitalen Fotos zu archivieren und sie so direkt einsteigen können und sich nicht durch verstreute allgemeine Informationen kämpfen müssen. Der in der Analyse gewonnene Eindruck, dass Formate und Formatmigrationen schwer zu

¹⁵³ NDIIPP 2010, S. 5

vermitteln sind, wird hier dadurch bestätigt, dass keine entsprechenden Hinweise formuliert sind und ausschließlich Bitstream Preservation mit jährlicher Prüfung der Lesbarkeit der Datei empfohlen wird. Dies stellt somit die einzige primäre Maßnahme dar, die in den Hinweisen nicht berücksichtigt wird.

Wünschenswert wären weiterführende Informationen auf der Website, z. B. zu Punkten wie Tagging (wie funktioniert das technisch, wie sollten die Begriffe gewählt werden). Hier könnte man dann auch Informationen zu Formatmigrationen und anderen aufwändigeren Themen bereitstellen.

5.1.4 Fazit

Der Vergleich zeigt, dass die drei Handreichungen zu Personal Digital Archiving auf unterschiedlichem Niveau und an unterschiedlichen Zielgruppen orientiert informieren. Die definierten primären Personal Digital Archiving Maßnahmen werden insbesondere von PARADIGM und Library of Congress/NDIIPP berücksichtigt, wohingegen die InterPARES-Hinweise nur wenige konkrete Hinweise zu den einzelnen Maßnahmen liefern:

	InterPARES	PARADIGM	LoC/NDIIPP
Dateiidentifikation	zu ausführlich	Sprechende Dateinamen, Ordnerstruktur, Metadaten	Sprechende Dateinamen, Metadaten, Ordnerstruktur, Index
Auswahl/ Bewertung	keine Angabe	“delete what’s not important”	“Pick the images you feel are especially important”
Formate	Hinweis auf offene Formate	ausführlich, mit konkreten Empfehlungen für die gängigsten Objekttypen	keine Angabe
Speichermethoden	Backup; keine Erläuterungen	Backup: SDD, entfernter Ort, Cloudspeicher	Backup: SDD, entfernter Ort, Cloudspeicher
Hardwareupdates	keine konkreten Hinweise	Enthalten	alle fünf Jahre

Tabelle 3: Berücksichtigung der primären Personal-Digital-Archiving-Maßnahmen in den einzelnen Anleitungen

Das Niveau der Guidelines des InterPARES2-Projekts liegt sehr hoch, da viele Details und Hintergrundinformationen gegeben werden und auch technische Kompetenzen der Anwender vorausgesetzt werden. Außerdem ist bei vielen Tipps der Aufwand der Umsetzung sehr hoch, was private Anwender abschrecken könnte. Die Zielgruppe ist nicht klar definiert und die Tipps werden eher aus der Sicht von Gedächtnisorganisationen formuliert als aus der Sicht von Anwendern mit unterschiedlichem technischem und fachlichem Hintergrund. Dies könnte darauf zurückzuführen sein, dass das InterPARES-Projekt generell aus der Institutionssicht angegangen wurde.

Ebenfalls aus der Sicht von Institutionen, die private Nachlässe und Archive z. B. von Politikern aufnehmen, wurde das PARADIGM-Projekt initiiert. Erfreulicherweise weisen die hieraus entstandenen Guidelines aber Hinweise auf Skalierungsmöglichkeiten auf, obwohl auch hier viele Details erläutert werden. Die einfachen Tipps in Satz-

form und die übersichtliche Tabelle machen es privaten Anwendern leicht, die für sie interessanten bzw. umsetzbaren Handlungen herauszusuchen und anzuwenden.

Die Personal Digital Archiving Broschüre der Library of Congress/NDIIPP verzichtet auf längeren Fließtext und Details. In Aufzählungsform werden für die häufigsten Typen digitaler Objekte die jeweiligen Handlungsschritte aufgelistet und in einfachen Sätzen in Befehlsform formuliert, so dass der Anwender mit der jeweiligen Anwendung zur Hand z. B. direkt mit der Archivierung seiner Fotos beginnen könnte. Hintergrundinformationen zu Datenmigration etc. werden nicht gegeben.

Eine deutschsprachige Aufbereitung von Hinweisen zu Personal Digital Archiving sollte sich an den Handreichungen des PARADIGM-Projekts und der NDIIPP orientieren, die sowohl die Zielgruppe Privatanwender ansprechen als auch ihre heterogenen technischen und fachlichen Kompetenzen berücksichtigen. Für die schnelle, handliche Information können die Flyer-artigen Handreichungen der NDIIPP als Grundlage dienen. Diese können sowohl als PDF auf einer Website bereitgestellt werden, als auch ausgedruckt als Flyer in Bibliotheken und anderen Einrichtungen ausgelegt werden. Ausführlichere Informationen könnten dann auf einer Website als Fließtext bereitgestellt werden, wobei die stichpunktartigen Tipps aus den Flyern dort ebenfalls wiederzufinden sein sollten, wie es auch bei den PARADIGM-Hinweisen der Fall ist.

5.2 Personal Digital Archiving als Teil der Informationskompetenzvermittlung an wissenschaftlichen Bibliotheken

Eine weitere Idee zur Vermittlung und Schulung von Personal Digital Archiving speziell für die Zielgruppe der Nutzer wissenschaftlicher Bibliotheken ist die Einbindung in Informationskompetenzschulungen.

Forschungsdatenmanagement und die Archivierung anderer wissenschaftlicher Primär- und Sekundärdaten in Hochschulrepositorien ist mittlerweile eine eigene Disziplin innerhalb der Langzeitarchivierungsforschung, zu der bereits umfangreiche Studien, Theorien und Best-Practice-Empfehlungen vorliegen. Die Archivierung wissenschaftlicher Daten von Forschern beginnt aber nicht erst mit dem Berufseinstieg. Bereits im Studium werden Daten in Form von Aufsätzen und Zeitschriftenartikeln recherchiert, gesammelt und genutzt sowie eigene Daten wie Seminar- und Abschlussarbeiten erzeugt und gesammelt. Neben den Daten mit Bezug zum Studium fallen natürlich bei allen Studierenden auch die üblichen persönlichen Daten wie Fotos, E-Mails etc. an, die gemeinsam eine private Kollektion bilden. Die Beschäftigung mit dem Thema Personal Digital Archiving könnte also schon zu diesem Zeitpunkt einsetzen, gesteuert durch entsprechende Angebote der wissenschaftlichen Bibliotheken.

Ellyssa Stern Cahoy beschäftigt sich in ihrer Forschung mit dem Thema Informationskompetenz und Personal Digital Archiving. Als Vorsitzende¹⁵⁴ des Ausschusses, der sich um die Neufassung der ACRL-Richtlinien für die Informationskompetenzvermittlung in den USA kümmert, regte sie eine Aufnahme von Personal Digital Archiving als Teil der Informationskompetenz an, was mit der erwarteten neuen Fassung auch umgesetzt sein soll.¹⁵⁵ Nach Ansicht von Cahoy eignen sich gerade die angebotenen Schulungen zu Zitierprogrammen wie Zotero, citavi oder Endnote für die Vermittlung von Personal Digital Archiving Themen, da bei der Nutzung dieser Programme Wissen und Objekte verwaltet, bewertet und gespeichert werden und somit diese auch für Personal Digital Archiving wichtigen Denkansätze aktiviert sind:

¹⁵⁴ Vgl. Cahoy 2013a, S. 12

¹⁵⁵ Vgl. Cahoy 2013b, S. 149

Every time a student or faculty member learns how to use citation management software, there is an opportunity to embed instruction on personal information management and preservation. With Zotero, Mendeley, and EndNote, libraries save and organize only one type of information—scholarly resources—but they present a structure to help users think more broadly about their information collections, how (and where) they are accessed, as well as where they are saved (and archived).¹⁵⁶

Cahoy sieht die Informationskompetenzvermittlung insbesondere unter dem Aspekt des *affective learning*, also dem Lernen über das Umgehen mit bestimmten Emotionen wie z. B. Angst, Frust, Enttäuschung innerhalb des Prozesses des Verfassens einer wissenschaftlichen Arbeit.¹⁵⁷ Daraus leitet sie ab, dass Personal Digital Archiving nicht von oben herab gelehrt, sondern in den individuellen Arbeitsablauf des Studierenden eingebaut werden sollte, so wie er damit am besten arbeiten kann.¹⁵⁸

Cahoy stellt aber auch heraus, dass die Vermittlung von Kompetenzen im Bereich Personal Digital Archiving ebenfalls eine Schulung der vermittelnden Bibliothekare erfordert, damit diese insbesondere technisch auf demselben Stand sind wie ihre Nutzer.¹⁵⁹ In der hochschulpolitischen Dimension sieht sie die Fachreferenten in der Pflicht, auch auf Seiten der Fakultäten entsprechende Aufmerksamkeit auf das Thema zu lenken, um die Integration in die Vermittlung zu unterstützen und die spätere Nutzung von Hochschulrepositorien zu verbessern.¹⁶⁰

Die Anwendbarkeit dieses Ansatzes in Deutschland müsste unter Einbeziehung der hier geltenden Rahmenbedingungen untersucht werden. Obwohl auch hierzulande die ACRL-Richtlinien als Grundlage anerkannt sind, ist Informationskompetenzvermittlung in Deutschland noch nicht auf demselben Stand wie in den USA, wo die *teaching library* bereits eine feste Größe ist. Untersucht werden müsste insbesondere, welche Resonanz Veranstaltungen zur Informationskompetenz hierzulande überhaupt haben, wie die vermittelnden Bibliothekare hinsichtlich Langzeitarchivierung geschult sind und inwiefern das Thema überhaupt unter dem Eindruck begrenzter Kontingente von Semesterwochenstunden für die Informationskompetenzvermittlung sinnvoll vermittelt werden kann oder ob Zusatzangebote über Lernplattformen in Form von freiwilligen Kursen angeboten werden können.

5.3 Personal Digital Archiving an öffentlichen Bibliotheken

Öffentliche Bibliotheken wie Stadt- und Stadtteilbibliotheken bieten schon seit langem neben der Möglichkeit der Ausleihe u.a. von Ratgeberliteratur, Belletristik und Kinder- und Jugendliteratur speziell auf ihre Nutzergemeinschaft zugeschnittene Dienstleistungsangebote. Dazu gehören unter anderem Computerkurse, Integrationsangebote und Hilfe bei der Nutzung sozialer Netzwerke wie Facebook. Da diese Bibliotheken den direkten Kontakt zu Menschen haben, die privat Daten erstellen, sammeln und verwalten, scheinen sie prädestiniert für die Vermittlung von Wissen zu Personal Digital Archiving. Aktuell bieten die drei größten Stadtbibliotheken¹⁶¹ München, Köln und Frankfurt aber keine Veranstaltungen zum Thema Personal Digital Archiving an. Wie bereits

¹⁵⁶ Cahoy 2013c, S. 146

¹⁵⁷ Vgl. Cahoy 2013b, S. 147

¹⁵⁸ Vgl. Cahoy 2013b, S. 148f.

¹⁵⁹ Vgl. Cahoy 2013d, S. 76

¹⁶⁰ Vgl. Cahoy 2013c, S. 149

¹⁶¹ Laut DBS 2012 sind München, Köln und Frankfurt die nach Einwohnerzahl drei größten öffentlichen Bibliotheken in öffentlicher Trägerschaft, siehe Anhang.

in Kapitel 2.7 erwähnt, werden in den USA bereits Thementage zu Personal Digital Archiving mit Hilfe der von der Library of Congress/NDIIPP bereitgestellten Materialien veranstaltet. Auch die Information mittels Broschüren oder Flyern, wie in Kapitel 5.1 angeregt, könnte Interessierten erste Anhaltspunkte liefern und generell Interesse für das Thema wecken.

Einen ambitionierten Ansatz stellen Andrea Copeland und Deborah Barreau vor: Sie wollen die Bibliotheksnutzer durch die Mitarbeit an einem öffentlichen Repository, das die Geschichte und das Geschichtserleben der jeweiligen Gemeinschaft dokumentiert, direkt in die verschiedenen Archivierungsprozesse einbeziehen und dadurch die Beschäftigung mit dem Thema und die Anwendung auf eigene private Kollektionen digitaler Objekte anregen:

We suggest that through co-created repositories, public library users can begin to understand the implications of their actions and of the systems they are using. With this understanding, they can effectively provide for long-term access to their digital information.¹⁶²

Ein wenig erinnert dieser Ansatz an das bereits vorgestellte *Digit*-Projekt des WDR: Man regt eine breite Zielgruppe, in jenem Fall die Zuschauer des WDR-Fernsehens, dazu an, ganz normale alltagsdokumentierende Objekte wie Fotos und Super8-Filme zu digitalisieren bzw. durch den WDR digitalisieren zu lassen, um ein Archiv entstehen zu lassen, das ein gewisses Gemeinschafts- und Identitätsgefühl vermittelt. Im Gegensatz zum *Digit*-Ansatz ist hier aber eine Institution Initiator, die als vertrauenswürdig wahrgenommen wird und keine versteckte Nutzung der Objekte für eigene Zwecke anstrebt; außerdem soll hierbei auch eine Diskussion entstehen zwischen den Mitgliedern der Gemeinschaft, die Objekte an das Repository weitergeben, Informationswissenschaftlern, die über das benötigte Wissen verfügen und den Mitgliedern der Gemeinschaft, die die angebotenen Objekte nutzen, kommentieren und dadurch angeregt ebenfalls Objekte zum Repository beisteuern.¹⁶³ Dadurch beschäftigen sich die Beitragenden automatisch mit verschiedenen Aspekten innerhalb des Archivierungsprozesses wie Datenschutz, Urheberrecht, Auswahl, Bewertung, Format und Speichermedien.¹⁶⁴ Eine Auseinandersetzung mit dem eigenen Archivierungsverhalten und den bekannten Unzulänglichkeiten könnte dadurch angeregt und damit eine Grundlage für entsprechende Schulungsaktivitäten geschaffen werden, z. B. hinsichtlich der Bewertung von Daten:

If individuals are encouraged to identify digital information of value to them, LIS professionals could then begin to teach public library users how to identify and preserve digital information of value to them through established curatorial practices.¹⁶⁵

Die öffentlichen Bibliotheken könnten sich nach Ansicht von Copeland und Barreau hierdurch innerhalb der Gemeinschaft neu aufstellen, indem sie zu ihrer alten sammelnden Rolle zurückkehren, um das kulturelle Erbe einer Gemeinschaft zu bewahren und in dieser Rolle auch wieder Menschen, die sonst keine Möglichkeit haben, sich entsprechendes Wissen anzueignen, bei Personal Digital Archiving unterstützen können.¹⁶⁶ Hierdurch könnte der Ansatz, steuergeldfinanzierte Forschung für den Normalverbrau-

¹⁶² Copeland und Barreau 2011, S. 642

¹⁶³ Vgl. Copeland und Barreau 2011, S. 639f.

¹⁶⁴ Vgl. Copeland und Barreau 2011, S. 639

¹⁶⁵ Copeland und Barreau 2011, S. 643

¹⁶⁶ Vgl. Copeland und Barreau 2011, S. 645

cher nutzbar zu machen, umgesetzt werden, ohne dass dies als Bevormundung empfunden wird und gleichzeitig die Bibliotheks- und Informationswissenschaft als gesellschaftlich relevante und nützliche Disziplin re-etabliert werden:

By helping individuals identify digital objects that are important to them and by increasing awareness of the technological challenges associated with digital object management, librarians and LIS researchers can become a driving force behind helping the public to preserve and organize their personal information collections.¹⁶⁷

Für eine Überprüfung der Umsetzbarkeit dieses Ansatzes in Deutschland müsste eine genauere Untersuchung des öffentlichen Bibliothekswesens, insbesondere der Verankerung der öffentlichen Bibliothek in der jeweiligen Gemeinschaft durchgeführt werden. Anzunehmen ist aber, dass z. B. in Städten mit gelebtem Lokalpatriotismus wie Köln ein entsprechendes Repositorium, z. B. unter dem Motto „Zeig uns Dein Köln – gestern und heute“ durchaus Anklang finden würde. Ein weiterer zu untersuchender Punkt ist natürlich die Finanzierung eines solchen Ansatzes, z. B. für den Betrieb des Repositoriums, Ausbildung der Bibliothekare, Veranstaltungen, Support etc.

¹⁶⁷ Copeland und Barreau 2011, S. 645

6 FAZIT

Personal Digital Archiving ist als Bereich der Langzeitarchivierung noch relativ wenig erforscht und in Deutschland nahezu unbekannt. Studien aus den Bereichen des Personal Information Management, der Archivwissenschaft und der Bibliothekswissenschaft beschreiben ein uneinheitliches Archivierungsverhalten, das durch unterschiedliche technische und organisatorische Fähigkeiten sowie divergierende Bewertung und Wertschätzung digitaler Objekte geprägt ist. Die Anknüpfungspunkte zur gut erforschten institutionellen Langzeitarchivierung sind relativ gering. Während auf institutioneller Ebene Archivierung ein emotionsloser Vorgang ist, der anhand festgeschriebener Regeln größtenteils automatisch durchgeführt wird, ist Personal Digital Archiving immer auch emotional geprägt, wodurch die Automatisierung erschwert wird. Aufgrund fehlender Softwarelösungen wurde die Schulung und Qualifizierung der privaten Anwender als bester Weg zur Verbesserung von Personal Digital Archiving in Deutschland ermittelt. Hierzu wurden primäre Maßnahmen analysiert, die auch in der institutionellen Langzeitarchivierung eingesetzt werden und in skaliert Form auf die private Archivierung angewendet werden können. Bibliotheken als Vertreter der institutionellen Langzeitarchivierung und als Informationsdienstleister mit direktem Kontakt zum Endanwender können diese Informationen verbreiten und schulen, unter Einbeziehung bestehender Kooperationen wie nestor. Drei verschiedene Vorschläge für die Umsetzung von Informations- und Schulungsangeboten wurden näher erläutert: Die Information der Allgemeinheit durch eine Website und Flyer, die sich an bestehenden englischsprachigen Angeboten orientieren könnten, unter Berücksichtigung der ermittelten primären Archivierungsmaßnahmen, die Integration von Personal Digital Archiving in Veranstaltungen zur Informationskompetenz an wissenschaftlichen Bibliotheken und die Vermittlung an öffentlichen Bibliotheken durch Aufbau eines gemeinschaftlichen Repositorien.

Bibliotheken könnten sich hierdurch als Bewahrer digitaler Information profilieren und Langzeitarchivierung generell für die Allgemeinheit verständlich machen. Dadurch könnte eine bessere Unterstützung von Langzeitarchivierungsprojekten, insbesondere unter finanziellen Gesichtspunkten erreicht werden. Auf der anderen Seite fließen so steuerfinanzierte Forschungsergebnisse zurück zum Steuerzahler und werden auch für ihn direkt nutzbar.

Hinsichtlich der Verbesserung der privaten Archivierung digitaler Objekte sollte aber immer beachtet werden, dass der private Nutzer nichts falsch macht; er handelt nach bestem Gewissen unter Einbeziehung seines vorhandenen Wissens und seiner technischen Möglichkeiten. Daher sollte die Motivation für die Erstellung von Informationsangeboten nicht darin bestehen, aus durchschnittlichen Computernutzern Archivare zu machen, sondern es muss immer das technische Umfeld und das vorhandene Wissen des Nutzers mit einbezogen werden. Dies bedeutet auch, dass die vermittelnden Bibliothekare ebenfalls sowohl in der Vermittlung geschult werden, als auch selber aktuelle technische Entwicklungen verfolgen, die sich auch auf das Archivierungsverhalten ihrer Nutzer auswirken.

Da die Studien von Marshall und Williams et al. beide mehr als fünf Jahre zurückliegen und somit Trends wie die zunehmende Nutzung von Cloud-Diensten, Smartphones und die Ablösung klassischer Softwareprogramme durch Apps auch auf Desktopcomputern noch nicht berücksichtigt werden konnten, besteht Bedarf an aktuellen Forschungsergebnissen. Studien mit deutschen Endanwendern könnten z. B. auch analysieren, ob deutsche Endanwender anders archivieren als die bisher untersuchten amerikanischen und englischen Endanwender und somit die bislang vorgefundenen Ergebnisse tatsächlich länderübergreifend für alle Nutzer generalisiert werden können.

Auch hinsichtlich der zu erstellenden Informationsangebote sollte weiter untersucht werden, welche anderen Informationskanäle sich insbesondere für die Vermittlung von Personal-Digital-Archiving-Techniken eignen, wie z. B. YouTube-Tutorials etc., mit denen der private Nutzer direkt erreicht werden kann.

Die in dieser Arbeit ausgelassenen Themen wie Datenschutz und Urheberrecht, insbesondere die Archivierung und Bearbeitung von Dateien, an denen nur ein Nutzungsrecht besteht, bedürfen einer separaten Untersuchung.

Für die Herstellung eines Problembewusstseins in Deutschland kann eventuell auch die Einigung auf einen entsprechenden Begriff von Vorteil sein, der selbsterklärender ist als Personal Digital Archiving, sich aber von anderen Begriffen wie Selbstarchivierung abgrenzt. nestor als deutsches Kompetenznetzwerk sollte hier aktiv werden und die Aufmerksamkeit, Qualifizierung und Information über seine Kooperationspartner und Kommunikationskanäle vorantreiben.

Hauptziel jeder Langzeitarchivierung, ob privat oder institutionell motiviert, sollte die Langzeitverfügbarkeit sein. Daher muss insbesondere der Bereich *Access*, wie der Zugriff und die Verwendung der archivierten Objekte im OAIS-Referenzmodell bezeichnet wird, beachtet und eventuell näher untersucht werden. Anders als in der institutionellen Langzeitarchivierung ist die Speicherung von Originalen und zur Nutzung bestimmten Derivaten bisher in der Personal-Digital-Archiving-Diskussion noch nicht explizit erörtert worden. Aber selbst wenn nur die Originale gespeichert bzw. migriert werden, sollte sichergestellt sein, dass der private Anwender diese Dateien weiterhin einfach öffnen, weiterbearbeiten und teilen kann. Denn aufwändig archivierte Daten, die am Ende aufgrund umständlicher Access-Methoden nicht genutzt werden, sind ebenfalls verlorene Daten.

7 QUELLEN- UND LITERATURVERZEICHNIS

Alle Onlinequellen zuletzt geprüft am 01.02.2015.

Beagrie, Neil (2005): Plenty of Room at the Bottom? Personal Digital Libraries and Collections. Online verfügbar unter <http://www.dlib.org/dlib/june05/beagrie/06beagrie.html>

Bell, Gordon; Gemmel, Jim (2007): Erinnerung total. In: *Spektrum der Wissenschaft* (05), S. 84-92. Online verfügbar unter <http://www.spektrum.de/alias/digitales-gedaechtnis/erinnerung-total/869379>

Brübach, Nils (2010): Das Referenzmodell OAIS. In: Neuroth, Heike et al. (Hg.): *nestor-Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung: Version 2.3*. Göttingen, S. 4:3-4:14.

Cahoy, Ellysa (2013a): The Quarterly Interview: Ellysa Stern Cahoy. In: *LOEX Quarterly* 40 (1), S. 12. Online verfügbar unter <http://commons.emich.edu/loexquarterly/vol40/iss1/5>

Cahoy, Ellysa Stern (2013b): Affective learning and personal information management. Essential components of information literacy. In: *Communications in information literacy* 7 (2), S. 146-149. Online verfügbar unter <http://www.comminfolit.org/index.php?journal=cil&page=article&op=view&path%5B%5D=v7i2p146&path%5B%5D=172>

Cahoy, Ellysa Stern (2013c): Faculty member as Archivist: Personal Archiving Practices in the Academic Environment. In: Donald T. Hawkins (Hg.): *Personal archiving. Preserving our digital heritage (Personal Information Management)*, S. 137-152.

Cahoy, Ellyssa Stern (2013d): The challenge of teaching personal archiving. In: *NDIIPP (Hg.): Perspectives on personal digital archiving*, S. 76-77.

CCSDS (2012): Reference Model for an Open Archival Information System (OAIS). Magenta Book. Online verfügbar unter <http://public.ccsds.org/publications/archive/650x0m2.pdf>

Copeland, Andrea J. (2011): Analysis of public library users' digital preservation practices. In: *J. Am. Soc. Inf. Sci.* 62 (7), S. 1288-1300. DOI: 10.1002/asi.21553.

Copeland, Andrea J.; Barreau, Deborah (2011): Helping People to Manage and Share Their Digital Information: A Role for Public Libraries. In: *Library Trends* 59 (4), S. 637-649. Online verfügbar unter www.ideals.illinois.edu/bitstream/handle/2142/26428/59.4.copeland.pdf

Cushing, Amber L. (2010): Highlighting the archives perspective in the personal digital archiving discussion. In: *Library Hi Tech* 28, S. 301-312.

DigCurV (2013): Skills in Digital Curation Curriculum Framework. Online verfügbar unter <http://www.digcurv.gla.ac.uk/>

Digital Curation Centre (Hg.): What is digital curation? Online verfügbar unter <http://www.dcc.ac.uk/digital-curation/what-digital-curation>

Engels, Melanie (2013): Vorbereitungen zur Langzeitarchivierung einer Fotokollektion. In: Achim Oßwald, Inka Tappenbeck, Haike Meinhardt und Hermann Rösch (Hg.): MALIS-Praxisprojekte 2013. Projektberichte aus dem berufsbegleitenden Masterstudiengang Bibliotheks- und Informationswissenschaften der Fachhochschule Köln. Wiesbaden: Dinges & Frick (B.I.T. online INNOVATIV, 44), S. 17-33. Online verfügbar unter http://www.b-i-t-online.de/daten/BIT_Innovativ_44_MaLIS.pdf

Ferle, Christoph H. (2012): Marktstudie digitale Langzeitarchivierung. Im Spannungsfeld zwischen Digital Preservation und Enterprise Information Archiving. Hg. v. Dieter Spath und Anette Weisbecker. Online verfügbar unter http://www.swm.iao.fraunhofer.de/content/dam/swm/de/documents/Marktstudie_Digitale_Langzeitarchivierung_web.pdf

Funk, Stefan E. (2010a): Kapitel 8.3: Migration. In: Neuroth, Heike et al. (Hg.): nestor-Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung: Version 2.3. Göttingen, S. 8:10-8:15.

Funk, Stefan E. (2010b): Kapitel 8.4: Emulation. In: Neuroth, Heike et al. (Hg.): nestor-Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung: Version 2.3. Göttingen, S. 8:16-8:23.

Gesetz über die Deutsche Nationalbibliothek (2006). Online verfügbar unter <http://www.gesetze-im-internet.de/bundesrecht/dnbg/gesamt.pdf>

GESIS (2013): Was wir sammeln (Collection Policy). Online verfügbar unter http://www.gesis.org/fileadmin/upload/institut/wiss_arbeitsbereiche/datenarchiv_analyse/2013-07-08_Was_wir_sammeln.pdf

GESIS (2014a): First Steps Towards Digital Preservation. Online verfügbar unter <http://www.gesis.org/veranstaltungen/workshops/2014-digital-preservation/>

GESIS (2014b): Introduction to Research Data Management for Social Scientists. Online verfügbar unter <http://www.gesis.org/veranstaltungen/workshops/2014-rdm/>

Grace, Stephen; Knight, Gareth; Montague, Lynne (2009): InSPECT Final Report. Online verfügbar unter <http://www.significantproperties.org.uk/inspect-finalreport.pdf>

Huppertz, Manfred (2013): Besser im Verbund - Kooperationen im Bereich der elektronischen Langzeitarchivierung. In: *Archivpflege in Westfalen-Lippe* (79), S. 19-21. Online verfügbar unter http://www.lwl.org/waa-download/archivpflege/heft79/19-21_huppertz.pdf

Huth, Karsten (2010): Kapitel 8.5 Computermuseum. In: Neuroth, Heike et al. (Hg.): nestor-Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung: Version 2.3. Göttingen, S. 8:24-8:31.

Huth, Karsten; Schroeder, Kathrin; Schumann, Natascha (2010): Kapitel 11.3 Speichersysteme mit Langzeitarchivierungsanspruch. In: Neuroth, Heike et al. (Hg.): nestor-Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung: Version 2.3. Göttingen, S. 11:7-11:14.

Indiana State University (Hg.): Personal Digital Archiving 2014. Online verfügbar unter <http://visions.indstate.edu/pda2014/index.html>

InterPARES 2 Project (Hg.) (o.J.): InterPARES 2 Project: products. Online verfügbar unter http://www.interpares.org/ip2/ip2_products.cfm

InterPARES 2 Project (2007): Creator Guidelines. Making and maintaining digital materials: Guidelines for individuals. Online verfügbar unter http://www.interpares.org/ip2/display_file.cfm?doc=ip2%28pub%29creator_guidelines_booklet.pdf

InterPARES Project (Hg.) (o.J.): InterPARES Project homepage. Online verfügbar unter <http://www.interpares.org/>

John, Jeremy Leighton; Rowlands, Ian; Williams, Pete; Dean, Katrina (2010): Digital Lives. Personal digital archives for the 21st century – an initial synthesis. Beta version 0.2. Online verfügbar unter <http://britishlibrary.typepad.co.uk/files/digital-lives-synthesis02-1.pdf>

JSTOR and the President and Fellows of Harvard College (Hg.) (2012a): JHOVE – JSTOR/Harvard Object Validation Environment. Online verfügbar unter <http://jhove.sourceforge.net/>

JSTOR and the President and Fellows of Harvard College (Hg.) (2012b): JHOVE Tutorial. Online verfügbar unter <http://jhove.sourceforge.net/using.html>

Keitel, Christian (2010): Kapitel 8.6 Mikroverfilmung. In: Neuroth, Heike et al. (Hg.): nestor-Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung: Version 2.3. Göttingen, S. 8:32-8:34.

Kim, Sarah (2013): Landscape of Personal Digital Archiving Activities and Research. In: Donald T. Hawkins (Hg.): Personal archiving. Preserving our digital heritage (Personal Information Management), S. 153-185.

Kulovits, Hannes; Becker, Christoph; Heister, Carmen; Rauber, Andreas (2010): Kapitel 13.2 Plato. In: Neuroth, Heike et al. (Hg.): nestor-Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung: Version 2.3. Göttingen, S. 13:3-13:20.

Library of Congress (Hg.) (o.J.a): Personal Digital Archiving. Online verfügbar unter <http://www.digitalpreservation.gov/personalarchiving/>

Library of Congress (Hg.) (o.J.b): Personal Digital Archiving Day Kit - Digital Preservation. Online verfügbar unter <http://www.digitalpreservation.gov/personalarchiving/padKit/index.html>

LOCKSS (Hg.) (o.J.): Preservation Principles | LOCKSS. Online verfügbar unter <http://www.lockss.org/about/principles/>

Marshall, Catherine C. (2007): How people manage information over a lifetime. Online verfügbar unter <http://www.csd1.tamu.edu/~marshall/PIM%20Chapter-Marshall.pdf>

Marshall, Catherine C. (2008a): Rethinking Personal Digital Archiving, Part 1: Four Challenges from the Field. In: *DLib Magazine* 14 (3/4). Online verfügbar unter <http://www.dlib.org/dlib/march08/marshall/03marshall-pt1.html>

Marshall, Catherine C. (2008b): Rethinking Personal Digital Archiving, Part 2: Implications for Services, Applications, and Institutions. In: *DLib Magazine* 14 (3/4). Online verfügbar unter <http://www.dlib.org/dlib/march08/marshall/03marshall-pt2.html>

Marshall, Catherine C. (2011): Challenges and Opportunities for Personal Digital Archiving. In: Lee, Christopher A., Society of American Archivists. (Hg.): I, digital. Personal collections in the digital era. Chicago: Society of American Archivists, S. 90-114. Online verfügbar unter <http://www.csdlib.tamu.edu/~marshall/I-Digital-Marshall.pdf>

Marshall, Catherine C. (2013): Social media, personal data, and reusing our digital legacy. In: Donald T. Hawkins (Hg.): Personal archiving. Preserving our digital heritage (Personal Information Management), S. 85-108.

Marshall, Catherine C.; Bly, Sara; Brun-Cottan, Françoise (2007): The Long Term Fate of Our Digital Belongings: Toward a Service Model for Personal Archives. Online verfügbar unter <http://arxiv.org/pdf/0704.3653>

Microsoft Research (Hg.) (2014): MyLifeBits – Microsoft Research. Online verfügbar unter <http://research.microsoft.com/en-us/projects/mylifebits/default.aspx>

NDIIPP (2010): Personal Archiving Brochure. Online verfügbar unter http://www.digitalpreservation.gov/personalarchiving/documents/PA_All_brochure.pdf

nestor (Hg.) (o.J.a): nestor-Wiki. LZA-Projekte und -Aktivitäten. Online verfügbar unter <https://wiki.dnb.de/pages/viewpage.action?pageId=94679349>

nestor (Hg.) (o.J.b): Startseite. Online verfügbar unter http://www.langzeitarchivierung.de/Subsites/nestor/DE/Home/home_node.html

nestor (Hg.) (2011): Memorandum of Understanding. Online verfügbar unter <http://nestor.sub.uni-goettingen.de/education/mou2011.pdf>

nestor-Arbeitsgruppe Policy (2014): Leitfaden zur Erstellung einer institutionellen Policy zur digitalen Langzeitarchivierung (nestor-Materialien, 18). Online verfügbar unter <http://nbn-resolving.de/urn:nbn:de:0008-2014052004>

Neuroth, Heike et al. (Hg.) (2010): nestor-Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung: Version 2.3. Göttingen. Online verfügbar unter http://www.nestor.sub.uni-goettingen.de/handbuch/nestor-handbuch_23.pdf

PARADIGM (Hg.) (2005a): paradigm | home. Online verfügbar unter <http://www.paradigm.ac.uk/index.html>

PARADIGM (Hg.) (2005b): paradigm | workbook. Online verfügbar unter <http://www.paradigm.ac.uk/workbook/index.html>

PARADIGM (2008a): Guidelines for creators of personal archives. Practical tips. Online verfügbar unter <http://www.paradigm.ac.uk/workbook/appendices/guidelines-tips.html>

PARADIGM (2008b): Guidelines for creators of personal archives. Eleven top tips for preserving your personal data. Online verfügbar unter <http://www.paradigm.ac.uk/workbook/appendices/guidelines-eleventips.html>

Peled, Ido (2011): The Challenges of Building Ex Libris Rosetta, a Digital Preservation System. In: *LIBER Quarterly; Vol 21, No 1 (2011)*. Online verfügbar unter <http://liber.library.uu.nl/index.php/lq/article/view/8012/8354>

PLANETS (Hg.) (2007): PLANETS homepage. Online verfügbar unter <http://www.planets-project.eu/>

Scheffel, Regine; Oßwald, Achim; Neuroth, Heike (2010): Kapitel 19: Qualifizierung im Themenbereich „Langzeitarchivierung digitaler Objekte“. In: Neuroth, Heike et al. (Hg.): nestor-Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung: Version 2.3. Göttingen, S. 19:1-19:20.

Schwens, Ute; Liegmann, Hans (2004): Langzeitarchivierung digitaler Ressourcen. Hg. v. Laisiepen, Klaus et al. München (Grundlagen der praktischen Information und Dokumentation, Bd. 1). Online verfügbar unter <http://files.d-nb.de/nestor/berichte/digitalewelt.pdf>

Strathmann, Stefan (2010): Rahmenbedingungen für die Langzeitarchivierung digitaler Objekte. In: Neuroth, Heike et al. (Hg.): nestor-Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung : Version 2.3. Göttingen, S. Kap. 3:1-Kap. 3:24.

Strodl, Stephan; Motlik, Florian; Stadler, Kevin.; Rauber, Andreas (2008): Personal & Soho Archiving. In: Proceedings of the 2008 Joint Conference on Digital Libraries. June 16-20, 2008, Pittsburgh, Pennsylvania. New York: ACM, S. 115-123.

Strodl, Stephan; Petrov, Petar; Greifeneder, Michael; Rauber, Andreas (2010): Automating Logical Preservation for Small Institutions with Hoppla. In: *Lecture Notes in Computer Science* 6273, S. 124-135.

The British Library (Hg.) (o.J.): Digital Lives – Main page. Unter Mitarbeit von Neil Beagrie. The British Library. Online verfügbar unter <http://www.webarchive.org.uk/wayback/archive/20120423072957/http://www.bl.uk/digital-lives/>

The British Library (Hg.) (2014): 404 - page not found on www.bl.uk. The British Library. Online verfügbar unter <http://www.bl.uk/digital-lives/>, zuletzt aktualisiert am 16.05.2014.

TU Wien (Hg.): Hoppla – Archiving System for Small Institutions. Unter Mitarbeit von Stephan Strodl. Online verfügbar unter <http://www.ifs.tuwien.ac.at/dp/hoppla/>

TU Wien (Hg.): Plato: The Preservation Planning Tool. Online verfügbar unter <http://www.ifs.tuwien.ac.at/dp/plato/intro/>

Vannevar Bush (1945): As We May Think. In: *The Atlantic* (7). Online verfügbar unter <http://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/>

WDR (Hg.) (2014a): Hilfe - WDR Digit. Online verfügbar unter <http://digit.wdr.de/howto>

WDR (Hg.) (2014b): Nutzungsbedingungen – WDR Digit. Online verfügbar unter http://digit.wdr.de/terms_of_use

WDR (Hg.) (2014c): Startseite – WDR Digit. Online verfügbar unter <http://digit.wdr.de/>.

Williams, Peter; Dean, Katrina; Rowlands, Ian; John, Jeremy Leighton (2008): Digital Lives. Report of Interviews With the Creators of Personal Digital Collections. In: *Ariadne* (55). Online verfügbar unter <http://www.ariadne.ac.uk/print/issue55/williams-et-al>

Williams, Peter; John, Jeremy Leighton; Rowland, Ian (2009): The personal curation of digital objects: A lifecycle approach. In: *Aslib Proceedings* 61 (4), S. 340-363. DOI: 10.1108/00012530910973767.

ANHANG

Größte öffentliche Bibliotheken in öffentlicher Trägerschaft nach Einwohnerzahl des Bibliotheksorts¹⁶⁸

Deutsche Bibliotheksstatistik (DBS) - Variable Auswertung

				BJ 2012
Nr.	Bibliothek	Name	DBS-ID	Nr. 1
				primäre Nutzer
	↑ ↓	↑ ↓	↑ ↓	↑ ↓
1	München StB	Münchner Stadtbibliothek	GE486	1.439.474
2	Köln StB	Stadtbibliothek Köln	AJ380	1.017.155
3	Frankfurt/M StB	Stadtbücherei, Zentralbibliothek	AA277	704.499
4	Stuttgart StB	Stadtbibliothek Stuttgart	AL480	616.137
5	Düsseldorf StB	Stadtbüchereien Düsseldorf, Zentralbibliothek	AE362	593.057
6	Dortmund StuLB	Stadt- und Landesbibliothek Dortmund	AA381	579.012
7	Essen StB	Stadtbibliothek	AE064	571.407
8	Bremen StB	Stadtbibliothek	AH478	547.340
9	Leipzig StB	Leipziger Städtische Bibliotheken	AJ197	531.809
10	Dresden StB	Städtische Bibliotheken Dresden	AE106	530.722
11	Leipzig BlindenB	Deutsche Zentralbücherei für Blinde	GE735	530.000
12	Hannover StB	Stadtbibliothek Hannover	AG115	527.135
13	Nürnberg StB	Stadtbibliothek im Bildungscampus Nürnberg	AB075	508.761
14	Duisburg StB	Stadtbibliothek	AJ136	486.752
15	Berlin Mitte StB	Stadtbibliothek Mitte, Hauptbibliothek Dr.-Philipp-Schaeffer	MB014	342.764
16	Berlin Tempelhof-Schönebg. StB	Stadtbibliothek Tempelhof-Schöneberg, Bezirkszentralbibliothek	GE036	334.965
17	Berlin Charlottenburg-W. StB	Stadtbibliothek Charlottenburg-Wilmersdorf	GC283	325.622
18	Berlin Neukölln StB	Stadtbibliothek Neukölln, Hauptbibliothek	GC301	320.667
19	Berlin Steglitz StB	Stadtbibliothek Steglitz-Zehlendorf	GE012	298.483
20	Berlin Friedrichsh.-Kreuzb.StB	Stadtbibliothek Friedrichshain-Kreuzberg (Berlin), Bezirkszentralbibliothek	MB038	276.396
21	Berlin Lichtenberg StB	Stadtbibliothek Berlin-Lichtenberg, Hauptbibliothek Anna Seghers	WB861	266.703
22	Berlin Marzahn ZB	Stadtbibliothek Marzahn-Hellersdorf Bezirkszentralbibliothek "Mark Twain"	WB782	254.381
23	Berlin Treptow-Köpenick StB	Stadtbibliothek Treptow-Köpenick	MB002	246.623
24	Berlin Reinickendorf StB	Stadtbibliothek Reinickendorf, Humboldt-Bibliothek	GD998	245.705
25	Berlin Spandau StB	Haupt- und Hauptjugendbibliothek Spandau	GD895	230.034
26	Berlin Krkhs	Patientenbibliothek CCM	MB920	38.072
27	Berlin Krkhs	Charité Campus Benjamin Franklin, Patientenbibliothek	GW108	38.041
28	München StB / Krkhs	Münchner Stadtbibliothek, Bibliothek im Klinikum rechts der Isar	KU157	N
29	München StB / Krkhs	Münchner Stadtbibliothek, Bibliothek im städtischen Krankenhaus Bogenhausen	KY250	N

Fragebogenfelder: 1. Einwohnerzahl des Bibliotheksortes

Bibliothekstyp: Öffentliche Bibliothek, Ergänzungsbibliothek, Patientenbibliothek, Musikbibliothek, Gefangenenbibliothek, Blindenbibliothek (einschränken)

Unterhaltsträger: Staat (Bund od. Republik), Land, Kreis/Bezirk, Kommune/Gemeinde, Sonstige öffentliche Trägerschaft, Körperschaft/Stiftung des öff. Rechts (einschränken)

Einwohnerzahl: über 400.000 (einschränken)

¹⁶⁸ Variable Auswertung für das Berichtsjahr 2012 durch <http://www.hbz-nrw.de/angebote/dbs/>